CRAS 2026

15th Conference on New Technologies for Computer and Robot-Assisted Surgery



16–18 SEPTEMBER 2026

University of Leeds · United Kingdom

Conference Proceedings

Surgical robotics · perception · autonomy · soft robotics · magnetics

UNIVERSITY OF LEEDS

ROBOTICS AT LEEDS

cras-eu.org

LEEDS · UK

15th Conference on New Technologies for Computer and Robot Assisted Surgery

CRAS 2026 — 16 to 18 September 2026

University of Leeds, UK

Hosted by University of Leeds, UK

https://cras-eu.org/cras-2026/

CRAS

Joint Workshop on Computer/Robot Assisted Surgery

UNIVERSITY OF LEEDS

Revision 3. Check for updates at https://cras-eu.org/cras-2026/abstracts/

This work is subjected to copyright. It is published as an open-access publication under the "Creative Commons Attribution 4.0 International" license.

Exhibitors and Sponsors

Hosted by:





Sponsors:









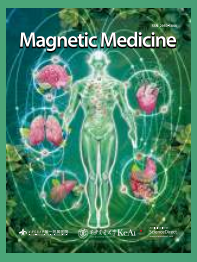






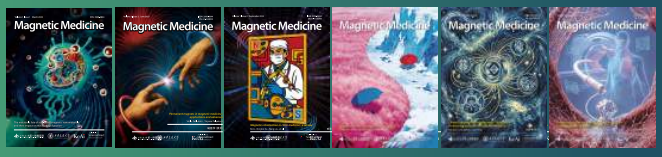
Magnetic Medicine

The first open-access journal dedicated to the biomedical and clinical applications of magnetic materials and technologies. (Quarterly, launced from 2025, gold OA with No APC).



·First Issue Launch: March 2025 | Quarterly Publication

Published on ScienceDirect - The world's leading scientific platform
100% Open Access - No copyright fees, maximum global visibility



Magnetic Medicine is proud to be a sponsor of CRAS2026. **A Special Issue focusing on magnetic medical robotics is currently being planned.** We warmly welcome submissions highlighting new insights into the **biomedical translation of magnetic-controlled and magnetically actuated medical robots and devices.** Accepted articles will be published with **NO APC**.

Article Highlights

- Towards an MRI-driven robotic cold-snare system for gastrointestinal polyp resection
- Magnetic robotization in clinic medicine: A review
- Magnetic steering for curved trajectories in neurosurgery — Applications and challenges in the brain and spinal cord
- Permanent magnets in magnetic medicine: Applications and advances

Journal Scope

Advancing Clinical Translation of Magnetic Technologies & Materials

01 Clinical Applications of Magnetic Technologies
02 Magnetically Actuated Medical Robotics & Devices
03 Biomedical Magnetic Materials & Nanomaterials
04 Magnetic Particle Imaging (MPI) & Magnetic Bioimaging

Scan & Connect: Access our official portal and become part of the world's leading magnetic medicine innovation community.
We warmly welcome your contribution and insight!

Editorial office: magmed@xjtu.edu.cn



Welcome

Ey up.

It is our great pleasure to welcome you to the 15th Conference on New Technologies for Computer and Robot-Assisted Surgery, CRAS 2026.

Over the next few days, researchers, clinicians, engineers, students and industry colleagues will share their latest work, discuss the challenges facing surgical care and, we hope, start some new collaborations along the way.

CRAS has always been a friendly and close-knit conference. Its informal atmosphere gives students and early-career researchers the chance to present their work, ask questions and talk openly with people from across the field. Whether this is your first CRAS or your fifteenth, we hope you will feel part of the community.

This year's programme reflects the breadth of current work in computer- and robot-assisted surgery. We are delighted to welcome our keynote and guest speakers, whose perspectives span medical technology and clinical practice.

The conference depends on the contributions of many people. We would like to thank our programme committee, reviewers, sponsors, volunteers, keynote and guest speakers and, above all, the authors whose work makes the programme possible.

We are very pleased to have you with us in Leeds for CRAS 2026, and we hope you enjoy the conference.

The CRAS 2026 Organising Committee

Organisation

Local Organising Committee

Pietro Valdastri, James Avery, James Chandler, Dominic Jones and Duygu Sarikaya

Programme Committee

James Avery, James Chandler, Leonardo De Mattos, Dominic Jones, Emmanuel Vander Poorten, Benoit Rosa and Duygu Sarikaya

Reviewers

Agostino Stilli
Albert Hernansanz
Ali Kafash Hoshiar
Alicia Casals
Antonia Tzemanaki
Benoit Rosa
Benoît Herman
Chao Liu
Christos Bergeles
Daniele Meli
Di Wu
Edward Grant
Emmanuel Vander Poorten
Evangelos B. Mazomenos
Ferdinando Rodriguez y Baena
Florent Nageotte
George Mylonas
Gernot Kronreif
Gianni Borghesan
Helge Wurdemann
Ilana Nisky
Iñigo Iturrate
Joao Bimbo
Kaspar Althoefer
Kathleen Denis
Leonardo De Mattos
Lin Cao
Lukas Lindenroth
M. Taha Chikhaoui
Manish Chauhan
Matthias Tummers
Nabil Zemiti
Nicola Piccinelli
Paolo Fiorini
Pierre Jannin
Pierre Renaud
Riccardo Muradore
S. M. Hadi Sadati
Shuhei Miyashita
Stamatia Giannarou
Tamas Haidegger
Thiusius Rajeeth Savarimuthu
Tim Horeman
Veronica Iacovacci
Zhuoqi Cheng

Local Organising Team

Clare Verity
Alison Whiteley
Katie McCarthy
Patricia Nicole Ramon

Conference Volunteers

Alexander Brown
Andrea Yanez Trujillo
Benjamin Calme
Charles Delorey
Jackson Williams
Yushi Guo
Nikita Greenidge
Peter Lloyd
Yael May
Jack Sherman
Jazmin Levy
Samwise Wilson

Contents

Exhibitors and Sponsors **ii**

Welcome **iv**

Organising Committee **v**

Contents **vii**

Keynote Speakers **viii**

Programme **xii**
Oral Sessions xii
Wednesday 16th xii
Thursday 17th xiii
Friday 18th xv
Poster Sessions xvii
Thursday 17th xvii
Friday 18th xix

Author Index **187**

Keynote Speakers



Prof. David Jayne

Professor of Surgery, University of Leeds
Honorary Consultant Surgeon, Leeds Teaching Hospitals NHS Trust

The Application and Challenges of Surgical Robotics in General Surgery

David is Bowel Cancer UK & RCS England Professor of Surgery at the University of Leeds and Honorary Consultant Surgeon at Leeds Teaching Hospitals NHS Trust. His clinical interests include robotic and minimally invasive surgery for colorectal cancer.

He is Chief Investigator for several NIHR portfolio clinical trials, including MRC/EME/NIHR ROLARR, HTA/NIHR FIAT, HTA/NIHR SaFaRI, and MRC/EME/NIHR IntAct. He is Director of the Leeds Institute for Medical Research, Clinical Director of the Leeds NIHR HealthTech Research Centre, and Clinical Director for the Leeds RCS England Surgical Trials Centre.

He sits on several funding panels, including NIHR Advanced Fellowships, NIHR i4i PDA, and NHSE/SBRI Cancer Innovations Awards. His translational research interests include the development and early-phase evaluation of surgical technologies.

Prof. Dan Stoyanov

Professor of Robot Vision, University College London
Chief Scientist for Digital Surgery at Medtronic



Translating Surgical AI Research into Medical Devices

Dan Stoyanov is a Professor at UCL Computer Science and Co-Director of the UCL Hawkes Institute. His research is focused on surgical AI and robotics, with a particular focus on computer vision in surgical endoscopy and minimally invasive surgery.

He is a Royal Academy of Engineering Chair in Emerging Technologies, in addition to leading large-scale projects funded by the Wellcome Trust and EPSRC. He is Chief Scientist at Digital Surgery, acquired by Medtronic plc, and co-founder of Odin Vision Ltd, acquired by Olympus, Panda Surgical Ltd, Enacuity Ltd, Helico Technologies Ltd, and Invexo Ltd.

He is also a Fellow of the Royal Academy of Engineering, the Academy of Medical Sciences, IEEE, IET, and the MICCAI Society.

Prof. Brad Nelson

Professor of Robotics and Intelligent Systems, ETH Zürich
Co-Founder, Nanoflex Robotics and Ophthorobotics



Clinically Ready Microrobots

Brad Nelson has been Professor of Robotics and Intelligent Systems at ETH Zürich since 2002, where his research focuses on micro-, nano-, and medical robotics. He studied mechanical engineering at Illinois and Minnesota, worked in computer vision at Honeywell and software engineering at Motorola, taught mathematics in Botswana, and then obtained a PhD in Robotics from Carnegie Mellon University.

At ETH Zürich, Nelson has built an internationally respected research group known for turning fundamental questions into working systems. His laboratory has become a training ground for researchers who value both rigorous science and practical impact, and many of his former students and collaborators now lead their own efforts in academia, industry, and medicine.

His work on microrobots has helped move the field from conceptual demonstrations towards clinical relevance.

Dr Massimiliano Simi

Executive Vice President of R&D,
Medical Microinstruments, Inc.



Robotic Open Microsurgery: Past, Present and Future

Dr Massimiliano "Max" Simi is Co-Founder and Chief Technology Officer of MMI, bringing over 15 years of leadership in surgical robotics. Driven by a passion for innovation, he invented the proprietary NanoWrist instrument, which forms the foundation of MMI's patent portfolio.

He continues to direct research and product development and upgrades for the flagship Symani Surgical System and Synaptix visualisation platform. Prior to co-founding MMI, Max served as an engineer, delivering advanced robotic and mechatronic solutions to Silicon Valley medical-device start-ups.

He holds a PhD in Innovative Technologies and Medical Robotics from Scuola Superiore Sant'Anna in Pisa, an Executive MBA, and both an MSc and BSc in Biomedical Engineering from the University of Pisa.

Prof. Stefanie Speidel

Professor in Translational Oncology,
National Center for Tumor Diseases Dresden



Beyond the Scalpel: From Intelligent Assistance to Autonomous Surgery

Stefanie Speidel has been Professor of Translational Surgical Oncology and a director at the National Center for Tumor Diseases Dresden since 2017. She is also one of the speakers of the DFG Cluster of Excellence CeTI and the Konrad Zuse AI School SECAI.

She received her PhD from Karlsruhe Institute of Technology in 2009 and led the junior research group Computer-Assisted Surgery from 2012 to 2016. She is Board Secretary and a Fellow of the MICCAI Society.

She has co-authored more than 170 publications, organises key workshops and challenges, and has served as general chair and programme chair for major international conferences, including IPCAI and MICCAI. Her research focuses on machine learning for image- and robot-assisted surgery.

Prof. Christian Duriez

Research Director, Inria
CEO, Compliance Robotics



From Surgical Simulation to Active Anatomical Models: Soft Robotics Applied to Surgery

Christian Duriez received a PhD in robotics from CEA and the University of Évry, France, in 2004. He was a postdoctoral fellow with the CIMIT SimGroup in Boston, Massachusetts.

In 2006, he joined Inria Lille Nord-Europe, where he worked on interactive simulation of deformable objects and haptic rendering, with a focus on surgical simulation. In 2014, he was promoted to Directeur de Recherche and is currently head of the DEFROST team, created in January 2015. In 2018, he was an invited researcher at Stanford University.

He contributed to the creation of the open-source SOFA framework and co-founded InSimo in 2013 and Compliance Robotics in 2024. His research interests include soft-robot modelling and control, fast finite-element methods, simulation of contact response, and the use of soft robots for haptics.

Honorary Guest Speaker

Prof. Hongliang Ren

Professor of Electronic Engineering,
Chinese University of Hong Kong



Prof. Hongliang Ren is a Professor in the Department of Electronic Engineering at the Chinese University of Hong Kong, where he received his PhD in 2008. His academic career has included appointments at Johns Hopkins University, Harvard Medical School, leading US children's hospitals, and the National University of Singapore.

His research interests include biorobotics, medical mechatronics, soft continuum robots, intelligent control, soft sensors, and multisensory learning for medical robotics.

He is a recipient of the National Science Fund for Distinguished Young Scholars, CUHK Young Researcher Award, NUS Young Investigator Award, IAMBE Early Career Award, and numerous best-paper awards.

He is consistently ranked among the world's top 2% most-cited scientists and has published more than 240 papers, receiving more than 23,000 citations, with an H-index of 77.

Programme

Oral Sessions

Wednesday 16th

Oral Presentations - 1 (Imaging & Visualisation) (15:00–16:00)

Session Chairs: Emmanuel Vander Poorten and Martina Finocchiaro

Title and authors	Page
15:00–15:10 — Advanced Anatomical Site Identification in Laryngoscopy: A Prototypical Contrastive Learning Approach for Quality Control Gaia Di Filippo, Aldo Marzullo, Claudio Sampieri, Elisa Bellini, Laura Ruiz Sevilla, Giorgio Peretti, Isabel Vilaseca, Leonardo S. De Mattos, Sara Moccia	1
15:10–15:20 — Online Guidewire Tracking in B-mode Ultrasound: A Kinematic-Gated U-Net Pipeline for Endovascular Robotic Guidance Chen Giladi, Lior Shilon	3
15:20–15:30 — A Synchronized Multimodal Data Collection Framework for AI-Based Autonomous Echocardiography Nicolai Jensen, Tobias Jacobsen, Thiusius R. Savarimuthu, Bruno Oliveira, Preman Kumarathurai, Axel Diederichsen, Viktor Poulsen	5
15:30–15:40 — EndoWrist Joint Configuration Estimation from Monocular Video via Depth Reconstruction and URDF Alignment Andrea Roberti, Riccardo Muradore, Nicola Massignani	7
15:40–15:50 — Optimizing Path Planning for Enhanced Robotic Ultrasound 3D Reconstruction Chiara Bernardini, Kaat Van Assche, Ayoob Davoodi, Marcello Bonfè, Emmanuel Vander Poorten, Gianni Borghesan	9
15:50–16:00 — Autonomous Endoscopic Navigation Based on Real-Time Lumen Segmentation Benjamin Calmé, Pietro Valdastri	11

Oral Presentations - 2 (Surgical Robots & Autonomy) (17:30–18:20)

Session Chairs: Arianna Menciassi and Thiusius Savarimuthu

Title and authors	Page
17:30–17:40 — **Remote Center of Motion Evaluation of a Compliant Joint for Neuroendoscopy** Federico Mariano, Elena De Momi, Giovanni Berselli, Leonardo S. Mattos	13
17:40–17:50 — **Computationally efficient workspace boundary determination for two-segment continuum robots** Joseph Massin, M. Taha Chikhaoui	15
17:50–18:00 — **Deformation-Aware Autonomous Circular Cutting in Robot-Assisted Surgery Using MFT-Based Tracking and Online Scissor Pose Planning** Yuriko Iyama, Pietro Valdastri, Dominic Jones	17
18:00–18:10 — **Delay Compensation for Micro-Macro Bilateral Teleoperation in Robot-Assisted Transoral Laser Microsurgery** Ayoob Davoodi, Sukrit Prasarnkleo, Mouloud Ourak, Gianni Borghesan, Emmanuel Vander Poorten	19
18:10–18:20 — **Inverse Design Framework for Tissue-Mimicking Phantoms: Applications in Computer and Robot-Assisted Surgery** Abu Bakar Dawood, Zhenyu Zhang, Dalia Osman, Cem Suulker, Arsen Abdulali, Martin Angelmahr, Kaspar Althoefer	21

Thursday 17th

Oral Presentations - 3 (Human Factors & Surgical Performance) (09:00–10:00)

Session Chairs: Ilana Misky and Leonardo De Mattos

Title and authors	Page
09:00–09:10 — **Bimanual Structure in Surgical Preparation** Eugenia Akpo, Nicholas Raison, Christos Bergeles, Letizia Gionfrida	23
09:10–09:20 — **Evaluation of Foundation Models for predicting visual outcomes after Epiretinal Membrane Surgery** Giulia Galvan, Ivo De Clerk, Mouloud Ourak, Elena De Momi, Emmanuel Vander Poorten	25
09:20–09:30 — **Vision-Based Hand-Motion Analysis for Ergonomic Assessment in Emulated Robot-Assisted Partial Nephrectomy** Deborah Cattafesta, Elisa Iovene, Mariachiara Di Cosmo, Federica Ferraguti, Sara Moccia	27
09:30–09:40 — **Towards Sensor-Free Surgical Skill Assessment Using Video-Based Multimodal Learning** Ali Osman, Duygu Sarikaya	29
09:40–09:50 — **Real-Time AI Generative Feedback for full-lung bronchoscopy instruction** Chuan Tian, Yu Hao, Haolin Jiang, Bruno Oliveira, Lars Konge, Kristoffer Cold, Zhuoqi Cheng	31

Title and authors	Page
09:50–10:00 — Real-Time Instance Segmentation in the Operating Room with Selective Refinement of Rare Surgical Instruments Rossella Damiano, Elena De Momi, Alberto Cesare Luigi Redaelli, Aldo Marzullo, Alice Ivanaj	33

Oral Presentations - 4 (Sensing & Control 1) (15:30–16:30)

Session Chairs: Di Wu and Nicola Picanelli

Title and authors	Page
15:30–15:40 — GelFlip: A Flip-Open Visual-Tactile Palpation Probe for Dual-Mode Sensing in Minimally Invasive Surgery Chuang Lu, Ziting Liang, Tianchi Chen, Yuyu Cai, Danail Stoyanov, Agostino Stilli	35
15:40–15:50 — Human-Robot Shared Control for Markerless Pedicle Subtraction Osteotomy Procedure Ingrid Tombini, Elena De Momi	37
15:50–16:00 — Electrical Impedance Tomography in Constrained Endoscopy Applications Chiara Falagario, Zhuoqi Cheng, Elena De Momi, Leonardo De Mattos	39
16:00–16:10 — Deep Learning-Based contact detection for Robot-Assisted retinal surgery using microscope and OCT imaging Giulia Galvan, Ivo De Clerck, Joy Haegemans, Mouloud Ourak, Emmanuel Vander Poorten	41
16:10–16:20 — Weakly Supervised Depth Estimation and Control of a Micropipette Huanyu Tian, Martin Huber, Lingyun Zeng, Zhe Han, Wayne Bennett, Giuseppe Silvestri, Gerardo Mendizabal-Ruiz, Tom Vercauteren, Alejandro Chavez-Badiola, Christos Bergeles	43
16:20–16:30 — A Tactile Sensing Array for Force Quantification in Robot Assisted Surgery Charles DeLorey, Michael Brockdorff, Adam Metcalf, Benjamin Calmé, Pietro Valdastri, Dominic Jones	45

Oral Presentations - 5 (Magnetics) (17:00–18:30)

Session Chairs: Veronica Iacovacci and Venkatasubramanian Kalpathy Venkiteswaran

Title and authors	Page
17:00–17:10 — An electromagnetic catheter for on-demand release and in-situ manipulation of magnetic therapeutic cargoes Riccardo Panizzolo, Tamerlan Srymbetov, Arianna Menciassi, Veronica Iacovacci	47
17:10–17:20 — Selectively Triggerable Biopsy Capsule via Multi-stage Magnetic Torsion Springs Simon Frieler, Jean-Paul P M de Vries, Sarthak Misra, Eric Diller, Venkatasubramanian Kalpathy Venkiteswaran	49

Title and authors	Page
17:20–17:30 — **Volumetric Magnetic Model and Polar-Coordinate Inverse Solver for Model-Based Control of a Magnetically Actuated Capsule Robot** Vito Romanello, Tamerlan Srymbetov, Arianna Menciassi, Veronica Iacovacci	51
17:30–17:40 — **Endoscopic Deployment of Magnetic Soft Continuum Robots for Targeted Gastrointestinal Therapy** Alistair Bacchetti, Alexander Brown, Benjamin Calmé, Pietro Valdastri, James H. Chandler	53
17:40–17:50 — **Forward-Model-Based Anomaly Detection for Magnetic Localization Systems** İsmail Şahin, Cengizhan Öztürk, Otar Akanyeti	55
17:50–18:00 — **Magnetic–Fluidic Soft Robotic Catheter Tip for Enhanced Access in Superior Mesenteric Artery Aneurysm Intervention** Muhammad Abdullah Khalid, Pietro Valdastri, James H. Chandler	57
18:00–18:10 — **A deployable, magnetically driven soft robotic platform for laparoscopic surgery: modeling, design and preliminary validation works** Matteo Bernabei, Lorenzo Mocellin, Matteo Cianchetti, Arianna Menciassi, Cesare Stefanini, Gastone Ciuti	59
18:10–18:20 — **AI Generated Synthetic Data for Accelerating the Development of Data-Driven Magnetic Localization Models** İsmail Şahin, Cengizhan Öztürk, Otar Akanyeti	61
18:20–18:30 — **Bi-directional Magnetic Vine Robot for Enhanced Endoscopic Navigation** Alexander Brown, Alistair Bacchetti, Benjamin Calmé, Joshua Davy, Sukjun Kim, Tania K Morimoto, James H. Chandler, Pietro Valdastri	63

Friday 18th

Oral Presentations - 6 (Novel Devices) (09:00–10:10)

Session Chairs: Agostino Stilli and Lin Cao

Title and authors	Page
09:00–09:10 — **A Low-Cost Smartphone-Based Diabetic Retinopathy Screening System Integrating AI Classification for Resource-Limited Settings** Ricardo Harrison, Gerardo Galindo, Nikita Greenidge, Sasha-Gay Wright, Pietro Valdastri	65
09:10–09:20 — **Magnetic Catheter Delivery with a Pressure-Lockable Hybrid Continuum Robot** Burak Ozdemir, Muhammad Abdullah Khalid, Benjamin Calmé, Pietro Valdastri, James H. Chandler	68

Title and authors	Page
09:20–09:30 — **A Miniaturized Everting Robot with a Large Working Channel for Ureteroscopy Using Twisted Tail Pushing** Juliana Dreyer, Cédric Girerd, Cécile Manceau, Nicolas Abdo, Nabil Zemiti	70
09:30–09:40 — **Towards Miniature Soft Vine Robots for Ultrasound Imaging of the Uterine Cervix** Andrea Libertad Yanez Trujillo, Nikita Greenidge, Benjamin Calmé, Joshua Davy, Pietro Valdastri, James H. Chandler, Virginia Pensabene, James Avery	72
09:40–09:50 — **Improved Passive Navigation in Soft Growing Eversion Robots Through Embodied Intelligence via Constrictive Bands** Cem Suulker, Muhie Al Haimus, Thomas Mack, Mohammad Sheikh Sofla, Neri Niccolo Dei, Reza Kashef, S.M.Hadi Sadati, Kaspar Althoefer	74
09:50–10:00 — **Origami Polyimide Sleeves for Protection of Variable Length Soft Endoscopic Actuators** Thomas Dean, Benjamin Calmé, Burak Ozdemir, Bao Nguyen, Pietro Valdastri, James H. Chandler	76
10:00–10:10 — **Towards Automated Fabrication of Soft Robotic Endoscope Tips for Affordable Gastroscopy** Alexander Brown, Pietro Valdastri, James H. Chandler	78

Oral Presentations - 7 (Sensing & Control 2) (16:00–17:00)

Session Chairs: Zhuoqi Cheng and Yao Zhang

Title and authors	Page
16:00–16:10 — **Towards a Sensorised Balloon for Lumen and Polyp Detection using Electrical Impedance Tomography** Jackson Williams, Pietro Valdastri, James Avery, James H. Chandler	80
16:10–16:20 — **Proximal Sensing of Instrument Rotation in Endovascular Procedures** Maria Lopes, Wim-Alexander Beckers, Emmanuel Vander Poorten, Beatriz Barata	82
16:20–16:30 — **Towards Combined Diffuse Reflectance and Electrical Impedance Spectroscopy for Colorectal Tissue Characterisation** Lukas Wachtel, Nilanjan Mandal, Jackson Williams, Andrea Libertad Yanez Trujillo, James Avery, Alexander J Thompson	84
16:30–16:40 — **Experimental Characterization of NURD in OCT Catheters Under Curvature and Rotational Speed Variations** Zeno Pogacnik, Giovanni Battista Regazzo, Beatriz Farola Barata, Daniel Corona Oliveira Costa, Emmanuel Vander Poorten, Oscar Caravaca-Mora	86
16:40–16:50 — **Shape control of a concentric tube continuum robot for surface ECoG brain implant insertion in a highly constrained environment** Ambroise Ulrich, Taha Chikhaoui, Blaise Yvert, Edwige Bano	88

Title and authors	Page
16:50–17:00 — **Magnetic Force Sensing in Ophthalmic Surgical Robot Forceps: Design, Integration, and Validation** Charles DeLorey, Benjamin Jackson, Mojtaba Esfandiari, Pietro Valdastri, Iulian Iordachita, Dominic Jones	90

Poster Sessions

Thursday 17th

Poster Pitches 1 (10:00–11:00)

Title and authors	Page
10:20–10:22 — **A Concentric Tube Catheter, Actuated and Imaged via MRI** Peter Lloyd, Yael L. May, Jurgen E. Schneider, Pietro Valdastri	92
10:22–10:24 — **Characterization of Robotic Needle Insertion for Automatic Biopsy Sampling of Dynamic Multilayer Polymeric Lung Phantoms** Maria Diaz, Gerardo Ramirez, Jorge Chairez	94
10:24–10:26 — **Teleoperation of a Six-DOF Parallel Manipulator for Spinal Decompression Surgery - Evaluation via Eggshell Grinding Experiments** Keng-Hua Yeh, Shih-Hsin Chen, Hsuan-Yu Chen	96
10:26–10:28 — **Compliance-Based Sensor Placement for Force Sensing on a Sensorized Prostate Phantom** Sizhe Tian, Yinoussa Adagolodjo, Jeremie Dequidt	98
10:28–10:30 — **Optimization-based compliance control for surgical continuum robots** Mattia Reina, Benoit Rosa, Hassan Omran	100
10:30–10:32 — **A Novel 2-DOF Continuous Hydraulic Motor for MR Intervention Robotics** Xinxin Li, Tse Zion	102
10:32–10:34 — **Elastic Loop Transfer Task: Bimanual Manipulation of Elastic Objects in Robot-Assisted Surgery** Shani Arusi, Uri Netz, Ilana Nisky	104
10:34–10:36 — **Towards minimally invasive in situ bioprinting using a soft endoscopic robot** Ruben Martin Rodriguez, Calvin Rüegg, Lorin Fasel, Nicolas Gerig, Valentina Basoli, Georg Rauter, Florian Thieringer	106
10:36–10:38 — **AR Visualization of Workspace Overlap for Collision Prevention in Robot-Assisted Surgery with Hugo™ RAS** Jonas Jakobi, Lukas Freiling, Do Yeon Kim	108
10:38–10:40 — **A Physics-Informed Neural Network for FBG-Based Force Vector and Location Identification along a Flexible Catheter** Xinran Liu, Emmanuel Vander Poorten, Mouloud Ourak, Johan Bennett	110

Title and authors	Page
10:40–10:42 — **Optimising Prostate Cancer Diagnosis with AI-Assisted Surface-Based Ultrasound-Derived Prostate-Specific Antigen Density** Rory Bennett, Tristan Barrett, Vincent Gnanapragasam, Zion Tse	112
10:42–10:44 — **Markerless Real-Time Target Tracking and Pose Estimation for Robot-Assisted Terahertz Imaging of Human Skin** Xiangyu Guo, Junlei Hu, Agrima Agarwal, Anubhav Dogra, Pietro Valdastri, Emma Pickwell-MacPherson	114
10:44–10:46 — **Applying AI techniques to EMG sensor data to predict localized muscle fatigue in RAS training: Standing and sitting postural approach** Daniel Caballero, Alfonso Fernández-Montero, Juan A. Sánchez-Margallo, Francisco M. Sánchez-Margallo	116
10:46–10:48 — **A Transparent Simulation-Centered Co-Robotic Platform for Human-Supervised Spinal Surgery** Daniel Shabanian, Pedro Machado, Kartikeya Walia, Philip Breedon	118
10:48–10:50 — **Development of a Pancreatic Phantom for Surgical Training** Arianna Callera, Francesco De Gaetano	120
10:50–10:52 — **Intuitive Robot-Assisted Drilling with Haptic Feedback and Multi-Layer Safety Constraints** Roman Leonov, Johannes Hölker, Do Yeon Kim, Dennis Kundrat, Jan E H Wolff	122
10:52–10:54 — **Kinematic considerations in laparoscopic surgical robots. Impact on usability** Josep Amat, Alicia Casals, Narcís Sayols	124
10:54–10:56 — **Physics-Guided Residual Modeling for Transient Prediction in Pneumatic Soft Continuum Robots** Wei Qiyu, Junbao Liu, Manish Chauhan	126
10:56–10:58 — **Quantitative Magnetic Centralisation for Continuum Robot MRI Localisation** Yael L. May, Peter Lloyd, James H. Chandler, Pietro Valdastri, Jurgen E. Schneider	128
10:58–11:00 — **2 Step Constraint Resolution With Regularization For Robotic Endoscopes With Environment Contacts** Maël Bianchi, Benoit Rosa, Hadrien Courtecuisse	130
11:00–11:02 — **Deep Learning 3D Surgical Tool Pose Estimation using Geometry-Aware Correction** Alexandros Dimitriadis, George Moustris, Costas Tzafestas	132
11:02–11:04 — **Exploiting Kinematic Redundancy for Active Stiffness Modulation of Continuum Concentric Push-Pull Robots** Azad Artinian, Taha Chikhaoui, Benoit Rosa	134
11:04–11:06 — **A SOFA-Based Virtual Reality Simulator for Vascular Procedures Integrated with the da Vinci Research Kit** Giulia Gamberini, Carlo Lodin, Selene Tognarelli, Arianna Menciassi	136

Friday 18th

Poster Pitches 2 (10:10–11:00)

Title and authors	Page
10:10–10:12 — Diffusion-Driven Safe Path Generation for Dual-Arm Surgical Autonomy Bita Azad, Siobhan Oca	138
10:12–10:14 — Visual servoing of continuum robotic endoscope for laser-based path scanning Xiao Xie, Gang Zheng, Isabelle Fournier, Nageotte Florent	140
10:14–10:16 — Magnetic Actuation of Vine Robots for Tortuous Surgical Navigations Joshua Davy, Nikita Greenidge, Nam Gyun Kim, Thomas Dean, Sukjun Kim, Shinwoo Park, Benjamin Calmé, Peter Lloyd, James H. Chandler, Jee Hwan Ryu, Tania K Morimoto, Pietro Valdastri	142
10:16–10:18 — Multi-fidelity model for thermal lesion prediction in robotic Ultrasound-Guided High-Intensity Focused Ultrasound surgery Francesca Parrotta, Giulia Gamberini, Selene Tognarelli, Giovanni Stabile, Arianna Menciassi	144
10:18–10:20 — Vision-Based Surgical Skill Assessment for Ring Transfer Tasks Using YOLO Segmentation Andrea Roberti, Sofia De Sanctis, Riccardo Muradore	146
10:20–10:22 — Advanced 5G Standalone Infrastructure for Remote Surgical Robotics: Design and Validation of the OSLO Platform Juan A. Sánchez-Margallo, Carlos Argent Masa, Francisco M. Sánchez-Margallo	148
10:22–10:24 — Evaluating localized muscle fatigue during suturing tasks in MIS surgical training Alfonso Fernández-Montero, Daniel Caballero, Manuel J. Pérez-Salazar, Juan A. Sánchez-Margallo, Francisco M. Sánchez-Margallo	150
10:24–10:26 — Preliminary Characterisation of Angular Overshoot and Oscillation in a Pneumatic Soft Continuum Robot Qiyu Wei, Haorui Sun, Manish Chauhan	152
10:26–10:28 — A Low-Power Multi-Sensor Navigation Architecture for Spatial Mapping of Endoscopic Capsules Sahil Anchal, Sean Coombes, Daragh Crowley, Eoin Higgins, Herman Alexander Jaeger, Pádraig Cantillon-Murphy	154
10:28–10:30 — Constraint Programming for Robotic Arm–Assisted Delivery of Ultrasound-Guided High-Intensity Focused Ultrasound (USgHIFU) Manisha Waterston, Lewis Howell, Bailey Andrew, James H. Chandler, Tze Min Wah, James McLaughlan	156
10:30–10:32 — Dense Estimation of Retinal Tectonics via Radial Basis Function Interpolation Tommaso Delucchi, Nicola Piccinelli, Andrea Roberti, Riccardo Mondin, Marco Pellegrini, Riccardo Muradore	158

Title and authors	Page
10:32–10:34 — **A Multimodal Laparoscopic Training Dataset for Automated Surgical Skill Assessment in Low-Resource Settings** Omar Choudhry, Sharib Ali, Chandra Shekhar Biyani, Dominic Jones	160
10:34–10:36 — **Benchmarking object detectors for naso-endoscopic anatomical features** Manuel Rocha, Willem Jansen, Cláudia Rocha, Armando Sousa, Maarten van Alphen, Michiel van den Brekel, Françoise Siepel	162
10:36–10:38 — **A Pre-Experimental Scan-Planning Simulator for Hexapod-Mounted On-Chip Terahertz Sensing** Nikita Greenidge, Connor D W Mosley, Josh P R Nixon, Cunningham John E., Pietro Valdastri	164
10:38–10:40 — **Trackerless Freehand Ultrasound 3D Reconstruction in Orthopedic Surgery** Jinghao Fei, Ruixuan Li, Carlos Rodriguez-Guerrero, Emmanuel Vander Poorten	166
10:40–10:42 — **Electrical Bio-Impedance Sensing on 2-Layered Tissue Model** Marianne Viola Petersen, Thiusius R. Savarimuthu, Zhuoqi Cheng	168
10:42–10:44 — **Feasibility of Machine Learning-Based Tissue Discrimination Using Electrical Bioimpedance Sensing** Yao Zhang, Anas El Hila, Teshome Getu Mhret, Mouloud Ourak, Emmanuel Vander Poorten	170
10:44–10:46 — **Evaluating Kinematic Tracking Stability and Load-Dependent Error in the da Vinci Research Kit (dVRK)** Christos Toilos, Evangelos Mazomenos, Laurent Mennillo	172
10:46–10:48 — **Pneumatically Actuated MRI-Compatible Robot for Spinal Needle Procedures** Yuxuan Qi, Aminul Ahmed, Christos Bergeles, Zion Tse	174
10:48–10:50 — **Vibroacoustic-Based Bone Density Differentiation During Robotic Drilling** Marthe Van den Bogaert, Maikel Timmermans, Kaat Van Assche, Ayoob Davoodi, Wim Desmet, Emmanuel Vander Poorten, Kathleen Denis	176
10:50–10:52 — **Design of Adjustable Constant-Force Surgical Forceps for Tissue Handling** Haopeng Lu, Zhuoqi Cheng, Thiusius R. Savarimuthu, Di Wu	178
10:52–10:54 — **An Open-Source Environment for Autonomous Robotic Colonoscopy Navigation** Guanglin Ji, Martina Finocchiaro, Kenny Erleben, Hang Yin	180
10:54–10:56 — **Intracardiac Soft Robotic Design for CPR Application** Zhihao Ma, Linxiaohai Ning, Xinxin Li, Tim Harris, Tse Zion	182
10:56–10:58 — **A Re-positionable and Axially Re-orientable Pivot-Joint Steering Mechanism for Tip-Everting Soft Growing Robots** Tianchen Ji, Diyang Wu, Zheyuan Bi, Lin Cao	184

Advanced Anatomical Site Identification in Laryngoscopy: A Prototypical Contrastive Learning Approach for Quality Control

Gaia Di Filippo[a,*], Aldo Marzullo[b], Claudio Sampieri[c,d], Elisa Bellini[d], Laura Ruiz Sevilla[c], Giorgio Peretti[e,f], Isabel Vilaseca[c], Leonardo S. Mattos[g], Sara Moccia[h]

[a]Department of Engineering and Geology, Università degli Studi "G. d'Annunzio" Chieti–Pescara, Italy
[b]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy
[c]Department of Otorhinolaryngology, Hospital Clínic de Barcelona, Barcelona, Spain
[d]Department of Medical Science (DIMES), University of Genoa, Genoa, Italy
[e]Unit of Otorhinolaryngology–Head and Neck Surgery, IRCCS Ospedale Policlinico San Martino, Genova, Italy
[f]Department of Surgical Sciences and Integrated Diagnostics (DISC), University of Genoa, Genova, Italy
[g]Biomedical Robotics Lab, Istituto Italiano di Tecnologia, Genova, Italy
[h]Department of Innovative Technologies in Medicine and Dentistry, Università degli Studi "G. d'Annunzio" Chieti–Pescara, Italy
[*]*gaia.difilippo@phd.unich.it*

INTRODUCTION

Early detection of laryngeal cancer is essential for minimizing patients' mortality and morbidity. Currently, white-light (WL) and narrow-band imaging (NBI) laryngoscopy represent the gold standard for optical biopsy in outpatient settings [1]. Laryngoscopy video examination depends highly on the operator's experience [1], and blind spots may occur during mucosal inspection, leading to incomplete coverage of the target areas. Kumai et al. [2] reported a missed diagnosis rate of approximately 85% during routine laryngoscopy in a population of 109 patients with hypopharyngeal cancer. Automated quality-control systems capable of assessing anatomical coverage may help reduce this rate, with potential benefits in terms of morbidity and mortality.

Recent literature has attempted to automate anatomical site identification using convolutional neural networks (CNNs), including Inception-ResNet-v2 [1] and EfficientNet-B0 [3]. However, fully supervised CNN-based approaches generally require large annotated datasets and may be sensitive to variations in acquisition setup, population characteristics, and image quality.

To address these limitations, we propose a pipeline that replaces direct pixel-based classification with a latent semantic representation of the anatomical regions observed during laryngoscopy. Rather than training a Vision Transformer (ViT) from scratch, we use a frozen self-supervised DINOv3 backbone to exploit visual representations learned from large-scale pretraining and train only a compact projection head. Supervised contrastive learning explicitly reduces intra-class variability while increasing inter-class separation, making the representation suitable for prototype-based classification under class imbalance and limited annotations. The resulting pipeline identifies anatomical sites according to their proximity to optimized synthetic prototypes. The method was evaluated on a dataset of 104 adult patients and 14 anatomical sites.

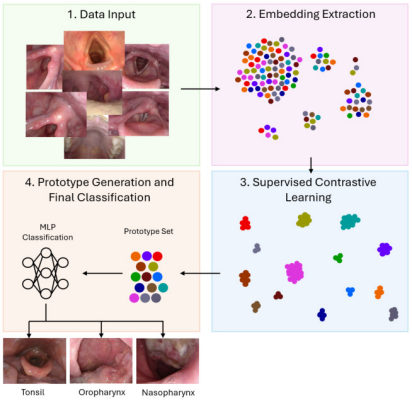


Fig. 1. Overview of the proposed anatomical site identification pipeline: informative-frame selection, frozen DINOv3 feature extraction, supervised contrastive projection, gradient-matched prototype generation, and final classification.

MATERIALS AND METHODS

The dataset used in this study was collected at the Hospital Clínic de Barcelona. A total of 104 videos were recorded during outpatient assessments with flexible fiberoptic laryngoscopy. Frames were automatically extracted from each video at a 1:10 rate and manually labeled into 14 anatomical classes by expert laryngologists. Figure 2 presents one representative frame for each class.

The dataset comprised 8,136 frames from 104 adult patients (> 18 years), including 51 healthy subjects and 53 patients with pathological conditions: squamous cell carcinoma, cysts, granulomas, leukoplakia, papillomas, polyps, and Reinke's edema. The class distribution was highly imbalanced, with frequencies ranging from 940 frames for the bilateral vocal folds to 77 frames for the nasal cavity.

The proposed pipeline (Fig. 1) consists of four phases: (i) *Data Input*, comprising informative frames. Following Baldini et al. [4], a frame was considered informative when it exhibited sufficient technical quality. A ResNet50 model pretrained on endoscopic images was used to discard non-

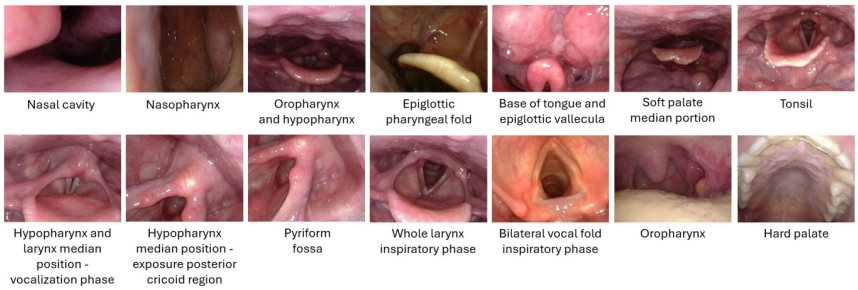


Fig. 2. Representative frames from the 14 annotated anatomical classes, illustrating the visual and anatomical variability encountered during laryngoscopy.

informative frames affected by blur, saliva, reflections, or underexposure; (ii) *Embedding Extraction*, in which the selected frames were mapped into a compact representation space using the frozen DINOv3 backbone; (iii) *Supervised Contrastive Learning*, in which embeddings were projected into a lower-dimensional latent space using a two-layer multilayer perceptron (MLP) with ReLU activation, batch normalization, and ℓ_2 normalization. The supervised contrastive loss encouraged embeddings from the same anatomical class to form compact clusters while separating different classes; and (iv) *Prototype Generation and Final Classification*, in which one synthetic prototype per class was initialized from a representative high-saliency embedding. The 14 prototypes were optimized through gradient matching over 1,000 iterations. A linear auxiliary classifier aligned the gradients obtained from the prototypes with those of the complete training set. Finally, a two-layer MLP was trained exclusively on the 14 optimized prototypes.

Unlike weighted cross-entropy or focal loss, which modify the contribution of individual samples during classifier optimization, the proposed contrastive stage acts directly on the geometry of the representation space. This choice was intended to obtain compact and inspectable class representations that can be summarized by one prototype per anatomical site.

RESULTS AND DISCUSSION

The pipeline was evaluated on a held-out test set comprising 20% of the patients, achieving an accuracy of 82%, a precision of 84%, and an F1-score of 82%. The confusion matrix in Fig. 3 shows the classification performance across the 14 anatomical sites. These results indicate that the proposed approach can provide useful discrimination despite the limited and imbalanced dataset.

For context, Li et al. [3] used an EfficientNet-B0 architecture to identify 11 anatomical sites from a dataset of 59,580 frames. Our pipeline considered 14 sites and was trained on approximately 5,000 frames, corresponding to about 8.4% of that number. However, this comparison is only contextual, since the two methods were evaluated on different datasets, populations, and class definitions. Therefore, the results suggest the potential data efficiency of the proposed representation, but do not establish superior performance or cross-dataset generalizability.



Fig. 3. Confusion matrix of the proposed model on the patient-level held-out test set. Rows represent ground-truth anatomical classes, while columns represent predicted classes.

CONCLUSION

This study presents a data-efficient pipeline for anatomical site identification in laryngoscopy based on self-supervised visual features, supervised contrastive learning, and optimized synthetic prototypes. The method achieved promising results on a limited and imbalanced dataset. Further ablation studies and external validation are required to quantify the contribution of each component and assess generalizability. The proposed approach could ultimately support laryngoscopy quality-control systems designed to identify incomplete anatomical coverage.

ACKNOWLEDGMENTS

This work was funded by the Horizon Europe project AIRCARE (GA no. 101137426).

REFERENCES

[1] J.-Q. Zhu, M.-L. Wang, Y. Li, W. Zhang, L.-J. Li, L. Liu, Y. Zhang, C.-J. Han, C.-W. Tie, S.-X. Wang *et al.*, "Convolutional neural network based anatomical site identification for laryngoscopy quality control: a multicenter study," *American Journal of Otolaryngology*, vol. 44, no. 2, p. 103695, 2023.

[2] Y. Kumai, T. Shono, K. Waki, D. Murakami, S. Miyamaru, Y. Sasaki, and Y. Orita, "Detection of hypopharyngeal cancer (tis, t1 and t2) by ENT physicians vs gastrointestinal endoscopists," *Auris Nasus Larynx*, vol. 47, no. 1, pp. 135–140, 2020.

[3] Y. Li, B. Ye, Y. Li, C. Fan, W. Chen, Y. Shuai, B. Liu, Q. Liu, K. Sun, W. Zhang *et al.*, "Real-time AI-assisted quality control during nasopharyngolaryngoscopy: a randomized controlled trial," *NPJ Digital Medicine*, 2026.

[4] C. Baldini, M. A. Azam, C. Sampieri, A. Ioppi, L. Ruiz-Sevilla, I. Vilaseca, B. Alegre, A. Tirrito, A. Pennacchi, G. Peretti *et al.*, "An automated approach for real-time informative frames classification in laryngeal endoscopy using deep learning," *European Archives of Oto-Rhino-Laryngology*, vol. 281, no. 8, pp. 4255–4264, 2024.

Online Guidewire Tracking in B-mode Ultrasound: A Kinematic-Gated U-Net Pipeline for Endovascular Robotic Guidance

Lior Shilon and Chen Giladi

Mechanical Engineering Department, SCE – Shamoon College of Engineering, Ashdod, Israel

chengi1@sce.ac.il (corresponding author)

INTRODUCTION

Endovascular interventions depend on continuous fluoroscopy, exposing patients and clinicians to cumulative ionizing radiation [1]. B-mode ultrasound is a radiation-free alternative, but reliable guidewire visualisation is hindered by speckle noise, anisotropic resolution, and the angle-dependent specular reflection of metallic wires [2], [3]. Although deep segmentation networks have recently demonstrated catheter detection in interventional ultrasound [4], two open problems block clinical deployment: **(i)** cross-anatomy generalisation across realistic vessel geometries, and **(ii)** robustness to operator-induced probe motion that corrupts the very segmentations the downstream geometry pipeline depends on. We address both with an integrated phantom-validated framework whose central novelty is a *kinematic gating* mechanism: an ultrasound-only classifier that gates the heavy segmentation engines frame-by-frame, holding the last valid tip estimate while the probe is detected as moving.

MATERIALS AND METHODS

Phantom and optical reference. A custom acrylic phantom houses four parallel channels at 5, 6, 7 and 8 mm with 2 mm TPU-90 acoustic linings (Fig. 1a). Fresh porcine tissue is overlaid above the phantom for realistic attenuation and clutter. The transducer carries an ArUco fiducial cube; an external camera, calibrated against a 25 mm chessboard target and time-synchronised to the ultrasound stream via MultiCam capture software, recovers a 6-DOF probe pose at 30 FPS that serves as both the localisation reference and the probe-stability label source. Acquisitions follow two protocols: *transducer-stationary* (probe held still, wire advanced/retracted) and *probe-motion* (wire held still, probe slid/lifted/re-contacted).

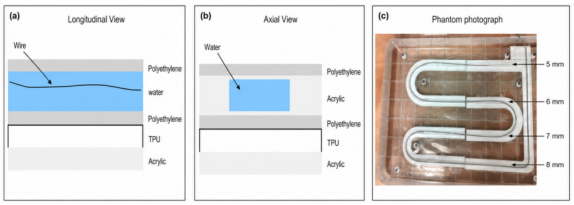
Two-stage segmentation and tip extraction. Two U-Nets [5] are trained on Phase I tissue+tube data (525 artery ROI masks; 375 wire masks) to produce binary masks M_t^A and M_t^W. The wire tip is the leftmost-column foreground in M_t^W (consistent proximal-to-distal insertion); a two-phase vertical intensity search rooted at the tip delivers upper- and lower-wall clearance.

Kinematic gating. For every frame we extract three feature streams *from the ultrasound feed*: global frame-difference, Farneback dense optical-flow magnitude (mean and variance), and eigenvalue variance / energy ratio of forward–backward dynamic mode decomposition over a 15-frame sliding window. Features are standardised via a 90-frame (∼3 s) rolling Z-score. The label comes from the optical reference: poses are Savitzky–Golay smoothed, then differentiated analytically through a 15-frame (±7) sliding 2nd-order polynomial fit; the gating label is the velocity-magnitude pair $(\|\mathbf{v}\|, \|\boldsymbol{\omega}\|)$. An XGBoost+SVM (RBF) ensemble is trained against this kinematic reference using 5-fold GroupKFold cross-validation grouped by source video sequence. A causal 5-frame median filter on the binary classifier output suppresses single-frame flicker at the cost of up to ∼133 ms decision delay at 30 FPS. The stability boundary ($v < 12.0$ mm/s, $\omega < 0.3$ rad/s) is determined empirically on a two-dimensional threshold grid.

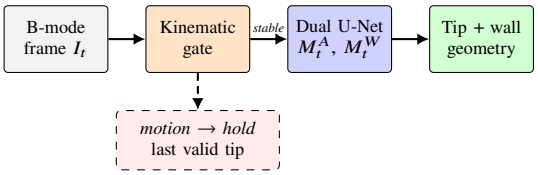
Cross-diameter generalisation. A 1,744-frame Phase II dataset, balanced at 436 image–mask pairs per diameter, drives a combinatorial leave-two-diameters-out protocol: six U-Net replicas are trained on every distinct diameter pair and tested on the held-out two.

RESULTS

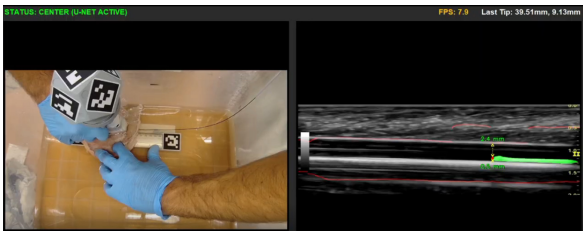
Segmentation. U-Net achieved Dice 0.965 (artery) and 0.942 (wire) on Phase I, vastly exceeding all classical thresholding/watershed baselines (Dice < 0.23 for wire). **Cross-diameter generalisation** ranged from Dice 0.751 (worst, train 6+7 mm → test 5+8 mm) to 0.819 (best, train 5+8 mm → test 6+7 mm), confirming that networks function as effective spatial *interpolators* but poor extrapolators (Fig. 1d). **Tip localisation** on the unseen 6 mm phantom (model trained on 7+8 mm) yielded mean signed displacement errors (per-frame predicted tip vs. optical reference) of −0.006 mm laterally and −0.033 mm axially with axial $\sigma = 0.188$ mm and $R^2 = 0.913$ (X), 0.961 (Y). A snapshot of the integrated live system—gating status, U-Net wire overlay, and on-screen tip coordinates in mm—is shown in Fig. 1c. **Continuous tracking** on the unseen 7 mm phantom (5+8 mm trained, *transducer-stationary* protocol with the wire actively advanced) maintained lateral $R^2 \approx 0.99$ and sub-millimeter axial mean SDE of 0.037 mm. Forced **geometric extrapolation** (5+6 → 8 mm) collapsed lateral R^2 to 0.106 and inflated lateral SDE to span −9.155 to +15.248 mm, motivating the operating-envelope boundary that gating respects. **Kinematic gating** reached SVM peak $F_1 = 0.781$; per-frame motion recall and false-negative rate at the deployed threshold (required for any safety claim) are reported in



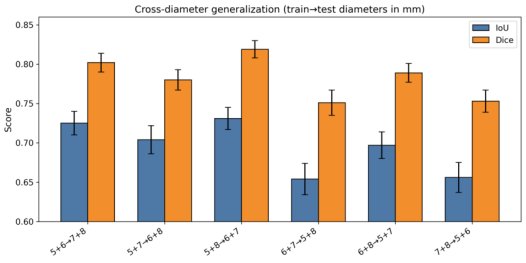
(a) Multi-diameter phantom: longitudinal cross-section, axial cross-section, and top photo of the four parallel channels (5, 6, 7, 8 mm).



(b) Pipeline: stable frames pass through the dual U-Net to tip + wall geometry; detected motion routes to a state-hold side branch.



(c) Live integrated system: operator view (left); U-Net wire overlay, gating status, mm tip coordinates, and FPS (right).



(d) Cross-diameter Dice / IoU across the six leave-two-out splits.

Fig. 1: **System and headline results.** (a) Acrylic multi-diameter phantom with porcine tissue overlay; the four S-shaped channels at 5/6/7/8 mm support cross-diameter generalisation tests. (b) Two-stage U-Net [5] backbone gated by an ultrasound-only stability classifier; motion routes to a state-hold side branch. (c) Live integrated system: operator with ArUco-tracked transducer (left, with the markers [6] providing 6-DOF reference via a calibrated external camera); ultrasound feed with U-Net wire overlay (green), gating status (`CENTER (UNET ACTIVE)`), tip coordinates in mm, and FPS counter (right). (d) Cross-diameter Dice / IoU sweep: networks trained on *extreme* diameters (5+8 mm) generalise best to intermediates (6+7 mm); training on adjacent mid-range pairs degrades generalisation.

the journal-version manuscript. Probe-motion sequences are handled at the system level via state-hold: when the gate fires, the last validated tip estimate is preserved until stability is re-detected. This is valid only while the wire is stationary during probe manipulation; concurrent wire advancement is enforced procedurally rather than detected, and producing a silent-drift indicator is required engineering before clinical use.

CONCLUSIONS AND DISCUSSION

Two operational principles emerge. First, U-Nets generalise cleanly *within* their trained geometric envelope but fail catastrophically outside it; bounding the training distribution by the maximum and minimum anatomical variance expected at deployment is necessary, and within this diameter-controlled phantom setting empirically effective [7] – a claim that must be re-examined for in-vivo curvature, branching, and calcification. Second, a deterministic geometric extractor (here, leftmost-pixel) running on a probabilistic segmentation has zero tolerance for fragmented masks; the current implementation does not apply connected-component filtering on M_t^W, so a single hallucinated foreground artifact triggers a centimeter-scale trajectory snap. Residual classifier errors and the brittleness of the extractor itself motivate a Kalman / HMM trajectory prior as future work. The gatekeeper is ultrasound-only at inference time, so the system is deployable on standard ultrasound hardware. Limitations: in-vitro phantom only and 2-D imaging; reproducing these results in vivo, on multiple ultrasound devices, and on 3-D matrix arrays remains the next stage of validation.

REFERENCES

[1] E. Vañó, L. González, J. M. Fernández, F. Alfonso, and C. Macaya, "Occupational radiation doses in interventional cardiology: a 15-year follow-up," *British Journal of Radiology*, vol. 79, no. 941, pp. 383–388, May 2006, doi: 10.1259/bjr/26829723.

[2] H. Sharei, T. Alderliesten, J. J. van den Dobbelsteen, and J. Dankelman, "Navigation of guidewires and catheters in the body during intervention procedures: a review of computer-based models," *Journal of Medical Imaging*, vol. 5, no. 1, p. 010902, Jan. 2018, doi: 10.1117/1.JMI.5.1.010902.

[3] A. Vandini, B. Glocker, M. Hamady, and G.-Z. Yang, "Robust guidewire tracking under large deformations combining segment-like features (SEGlets)," *Medical Image Analysis*, vol. 38, pp. 150–164, May 2017, doi: 10.1016/j.media.2017.02.001.

[4] A. Ranne, Y. Velikova, N. Navab, and F. Rodriguez y Baena, "AiAReSeg: Catheter detection and segmentation in interventional ultrasound using transformers," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 8187–8194, doi: 10.1109/ICRA57147.2024.10611539.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, ser. Lecture Notes in Computer Science, vol. 9351. Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[6] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, Jun. 2014, doi: 10.1016/j.patcog.2014.01.005.

[7] D. R. DeVries, L. J. Olafsen, J. S. Olafsen, H. H. Nguyen, K. E. Schubert, S. Dayawansa, and J. H. Huang, "Ultrasound localization of nitinol wire of sub-wavelength dimension," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 3, pp. 18–24, 2022, doi: 10.1109/OJEMB.2022.3151230.

A Synchronized Multimodal Data Collection Framework for AI-Based Autonomous Echocardiography

Nicolai H. Jensen[1], Tobias R. Jacobsen[1], Thiusius R. Savarimuthu[1], Preman Kumarathurai[2], Axel Diederichsen[2], Viktor Poulsen[2], and Bruno Oliveira[1]

[1]*SDU Robotics, Maersk McKinney Moller Institute, University of Southern Denmark, Odense, Denmark*
[2]*Department of Cardiology, Odense University Hospital, Odense, Denmark*

INTRODUCTION

Advances in social, economic, and healthcare development have reduced mortality, contributing to a globally ageing population. In 2021, chronic diseases accounted for at least 75% of non-pandemic-related deaths, with cardiovascular disease (CVD) the leading cause (~ 19 million deaths annually), with age as a strongly associated risk factor [1]. Medical imaging, such as ultrasound (US), is central to facilitate early cardiac diagnosis to improve treatment outcomes due to its safety and accessibility. However, it remains highly operator-dependent, requiring precise probe positioning and pressure, and is challenged by a projected global shortage of 11 million healthcare workers by 2030 [2]. This motivates the development of systems that reduce reliance on operator expertise.

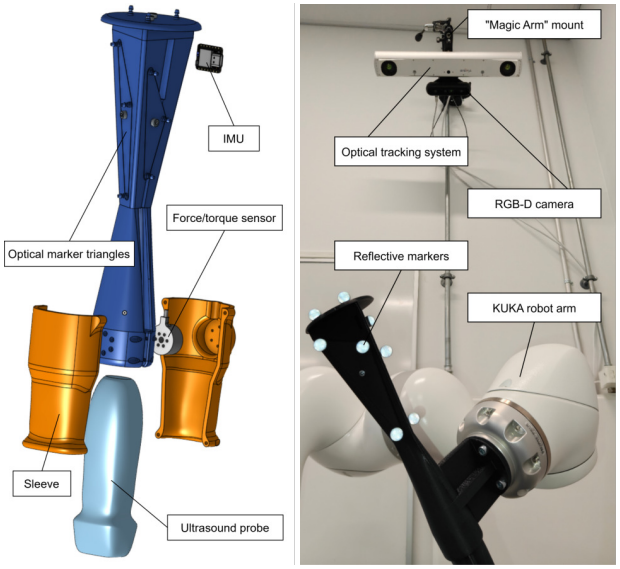
Artificial intelligence (AI) has shown strong potential in automating echocardiographic tasks, including view classification, segmentation, and diagnosis. The introduction of echo-specific foundation models promises to enhance trust and improve generalization, while vision-language models (VLMs) enable multimodal reasoning and human-robot interaction. Emerging world models further incorporate physical and spatial understanding of the environment [3]. Vision-language-action models (VLAs) represent a natural next step toward autonomous US acquisition but remain under-explored in the medical field [4]. Progress is limited by the scarcity of large, high-quality multimodal datasets required to capture expert behaviour and enable robust data fusion to better approximate the complexity of human perception [2].

Beyond US imaging, datasets for autonomous echocardiography should include probe position and orientation, which in current methods are obtained via IMUs, hand-guided robots, electromagnetic (EM) tracking, or optical tracking systems [3], [5]. Each approach presents trade-offs: IMUs are prone to noise, bias and drift; robot-assisted acquisition may alter clinical workflow; EM tracking is sensitive to interference; and optical tracking, while highly accurate, can suffer from marker occlusion.

This paper proposes a multimodal data collection framework for echocardiography integrating US imaging, probe pose, force/torque (F/T) sensing, RGB-D data, and inertial measurements (linear acceleration and angular velocity). The framework further incorporates methods for rigid body transformations and temporal synchronization.

MATERIALS AND METHODS

The sensors used to gain environment information in this data collection framework are the Atracsys fusionTrack™ 500 optical tracking system, for probe pose retrieval, and the Intel RealSense™ D555 RGB-D camera, for scene context. The manufactured probe structure is shown in Fig. 1a, which consists of a GE 4Vc-D US probe, a Wittenstein HEX21 six-axis F/T sensor, and a Seeed Studio XIAO nRF52840 board with an IMU. Four optical marker triangles with distinct characteristics and reflective spheres were designed and strategically positioned to ensure at least one is always visible to the tracking system.



(a) Exploded probe view. **(b)** Robot test setup.

Fig. 1: Sensors and components included in the US data collection framework.

The pose of the US probe tip is given by the series of rigid body transformations in Eq. 1, where B is the optical tracking base frame, M is an optical marker, R is a reference frame defined on the probe structure, and T is the tip. ${}^{B}_{M}T$ is given by the optical tracking system.

$$ {}^{B}_{T}T = {}^{B}_{M}T \; {}^{M}_{R}T \; {}^{R}_{T}T \tag{1} $$

The acceleration a measured by the IMU is transformed to the tip using Eq. 2.

$$ a_T = a_I + \dot{\omega} \times {}^{I}_{T}p + \omega \times \left(\omega \times {}^{I}_{T}p \right) \tag{2} $$

The angular velocity ω measured by the IMU is similarly transformed to the tip using Eq. 3.

$$\omega_T = {}^{T}_{I}R\ \omega_I \tag{3}$$

To transform the wrench w, as given by the F/T sensor, to the tip, the inverse transpose of the adjoint transformation is applied as seen in Eq. 4.

$$\mathrm{w}_T = Ad^{-T}_{{}^{F}_{T}T} \begin{bmatrix} \mathrm{f}_F \\ \tau_F \end{bmatrix} \tag{4}$$

All transformations T, containing R and p, are found in a similar way as in Eq. 1.

RESULTS

A KUKA LBR Med 7 robot arm, as seen in Fig. 1b, was used to perform experiments to validate the data collection framework, both in terms of accuracy of measurements and temporal synchronization. To determine the pose tracking accuracy, the robot was programmed to follow a path, at different velocities, where all optical markers would get tracked. A snippet of this trajectory is seen in Fig. 2.

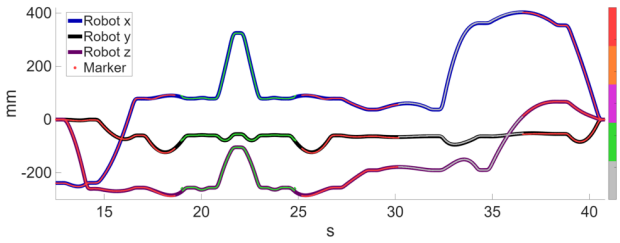


Fig. 2: Probe tip trajectory measured by robot and three different tracked markers (red, green, grey).

It can be seen that there are small discontinuities in the optical marker data when compared to the robotic path, especially at the transitions between markers, but still, the position and orientation errors are low, as seen in Fig. 3.

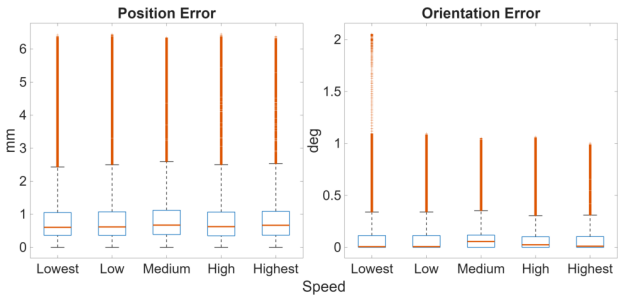


Fig. 3: Tracking accuracy of the optical markers, at different velocities, when compared to the robot pose.

The mean position tracking error, using the Euclidean norm, is 1.01 ± 1.23 mm, while the mean orientation tracking error, using the geodesic distance, is 0.11 ± 0.22°.

To compare the magnitude of force measured at the probe tip by the robot and by the force sensor, the robot was programmed to move up and down into contact with a phantom model. As seen in Fig. 4, the magnitudes are closely aligned, and the mean difference between robot and sensor is 1.41 ± 0.82 N.

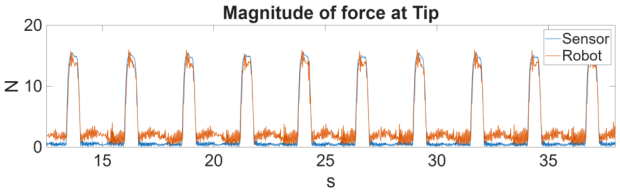


Fig. 4: Magnitude of force at probe tip measured by robot and force sensor.

Finally, the robot moved up and down while imaging the flat bottom of a water-filled container, creating periodic motion visible across multiple sensor modalities to determine the temporal alignment [6]. An impulse-based experiment, where the probe was tapped 40 times, was also conducted, as this more clearly revealed the delay for some modalities. The results can be seen in Table I.

TABLE I: Delay of different sensor modalities.

Modality	F/T	IMU	Optical	Robot	US image
Delay [ms]	*baseline*	1.0	6.7	35.4	43.8

CONCLUSIONS AND DISCUSSION

The results indicate that the collected data is of high quality, with a position tracking error of 1.01 ± 1.23 mm, and where the difference in force magnitude between robot and sensor (1.41 ± 0.82 N) can be largely attributed to the steady-state force measured by the robot. The average sensor delay can be compensated by subtracting the value from every timestamp, and the US imaging delay of 43.8 ms corresponds to ~ 2 frames at 50 FPS. Furthermore, this delay is similar to what other studies have found [6].

In conclusion, the accuracy of the proposed framework is acceptable for collection of high-quality multimodal data. The framework is already in use and being verified by cardiologists at OUH. Future work includes patient-mounted markers or sensors to further improve data quality by enabling motion compensation during acquisition, and accurate calibration of the US probe tip (origin of beam) with the optical tracking markers.

REFERENCES

[1] Y. E. Tian, V. Cropley, A. B. Maier, N. T. Lautenschlager, M. Breakspear, and A. Zalesky, "Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality," *Nature medicine*, vol. 29, no. 5, pp. 1221–1231, 2023.

[2] Z. Jiang, S. E. Salcudean, and N. Navab, "Robotic ultrasound imaging: State-of-the-art and future perspectives," *Medical Image Analysis*, vol. 89, p. 102878, Oct. 2023.

[3] Y. Yue, Y. Wang, H. Jiang, P. Liu, S. Song, and G. Huang, "EchoWorld: Learning Motion-Aware World Models for Echocardiography Probe Guidance," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2025, pp. 25 993–26 003, iSSN: 2575-7075.

[4] K. Kawaharazuka, J. Oh, J. Yamada, I. Posner, and Y. Zhu, "Vision-language-action models for robotics: A review towards real-world applications," *IEEE Access*, 2025.

[5] R. Droste, L. Drukker, A. T. Papageorghiou, and J. A. Noble, "Automatic Probe Movement Guidance for Freehand Obstetric Ultrasound," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Cham: Springer International Publishing, 2020, pp. 583–592.

[6] E. M. Moult, A. Lasso, T. Ungi, C. Pinter, M. Welch, and G. Fichtinger, "Improved temporal calibration of tracked ultrasound: An open-source solution," *Journal of Medical Robotics Research*, vol. 2, no. 04, p. 1750008, 2017.

EndoWrist Joint Configuration Estimation from Monocular Video via Depth Reconstruction and URDF Alignment

Nicola Massignani, Andrea Roberti, and Riccardo Muradore

Department of Engineering for Innovation Medicine, University of Verona, Strada le Grazie 15, 37134, Verona, Italy

INTRODUCTION

Robotic-assisted surgery platforms such as the *da Vinci* system generate large amounts of visual and kinematic data, enabling data-driven surgical analysis. However, in clinical settings robotic joint measurements are often unavailable, and datasets frequently consist only of monocular endoscopic RGB video, preventing accurate reconstruction of articulated instrument motion. This work addresses the problem of estimating *EndoWrist* joint configurations directly from monocular surgical recordings. Unlike learning-based joint estimation methods, which require large annotated datasets, or multi-view reconstruction approaches that depend on multiple synchronized cameras, the proposed framework combines monocular depth estimation with model-based geometric optimization and requires only a single RGB endoscopic video. We propose a vision-based pipeline that predicts depth using *Depth Anything v3*, reconstructs a 3D point cloud, and aligns it with a URDF-based EndoWrist model. Joint parameters are estimated by minimizing a robust geometric alignment error using ICP registration combined with *Particle Swarm Optimization (PSO)*. Because the registration objective is highly non-convex due to partial observations, self-occlusions, and cylindrical symmetries of the instrument shaft, a gradient-free optimization strategy is adopted to improve robustness against local minima. The proposed framework is evaluated on the *da Vinci Research Kit (dVRK)* using ground-truth robot kinematics for quantitative assessment. Experimental evaluation on the dVRK shows mean joint estimation errors of 0.073 rad (pitch), 0.099 rad (yaw), and 0.036 rad (gripper) using stereo point clouds, while monocular depth reconstruction achieves errors of 0.229 rad, 0.338 rad, and 0.059 rad, respectively, demonstrating the feasibility of recovering physically plausible EndoWrist joint configurations directly from monocular RGB video.

MATERIALS AND METHODS

The proposed framework estimates EndoWrist joint configurations through monocular depth reconstruction and model-based point cloud alignment. The pipeline consists of five stages: instrument segmentation, depth estimation, reference point cloud generation, ICP-based registration, and PSO-based joint optimization.

1) Instrument Segmentation: To isolate the surgical instrument from background tissue, we employ a YOLO-based instance segmentation network. Focusing on the dVRK Patient Side Manipulator (PSM1). This mask extracts the instrument region of interest, reducing background clutter and improving the robustness of subsequent 3D reconstruction.

2) Monocular Depth Estimation & Point Cloud Generation: Depth estimation is performed using Depth Anything v3 [1], which predicts a dense depth map from the segmented RGB image. Because monocular depth inherently suffers from scale ambiguity and only captures unoccluded surfaces, the resulting point cloud is metrically inconsistent with the URDF model and requires explicit scale normalization prior to registration.

3) Reference Point Cloud Generation: A synthetic point cloud is generated by uniformly sampling 1000 points per link from the EndoWrist URDF mesh. To match the partial visibility of the monocular reconstruction, we apply a *back-face culling* strategy, discarding faces whose outward normal $\mathbf{n}$ satisfies $\mathbf{n} \cdot \mathbf{z}_{\text{cam}} > 0$.

4) Point Cloud Registration & Alignment: To evaluate a candidate joint configuration, we rigidly align the observed cloud $\mathcal{P}_s$ and the URDF-derived cloud $\mathcal{P}_t$ using Iterative Closest Point (ICP). To address scale ambiguity, both point clouds are first mean-centered. A global isotropic scale factor α is then estimated by computing the ratio between the characteristic radial extents of the two point sets: $\alpha = \frac{r(\tilde{\mathcal{P}}_t)}{r(\tilde{\mathcal{P}}_s)}$, where $r(\cdot)$ represents a robust measure of the point cloud's radial distribution. This normalization ensures that the reconstructed geometry is metrically consistent with the URDF model prior to registration.

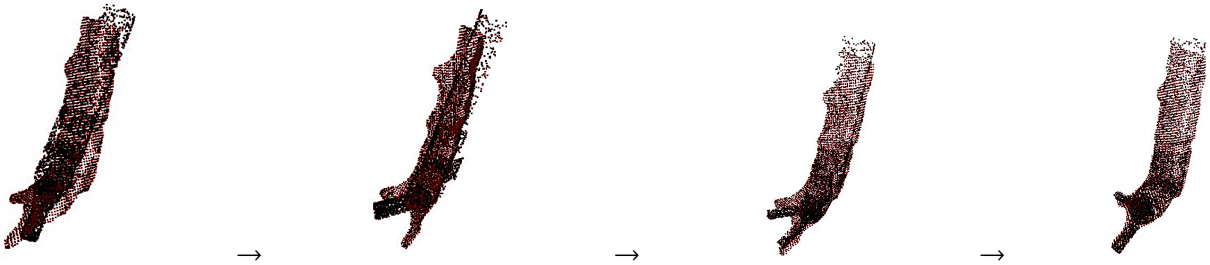
Due to the instrument's cylindrical symmetry, ICP is initialized with multiple longitudinal rotations to avoid local minima. The final alignment is evaluated using a robust symmetric fitness metric based on a truncated one-way error: $E_{\text{sym}}(\mathcal{P}_s, \mathcal{P}_t) = E(\mathcal{P}_s \to \mathcal{P}_t) + E(\mathcal{P}_t \to \mathcal{P}_s)$,

5) Joint Configuration Estimation: Joint estimation is formulated as a non-convex optimization task where the objective function is the robust registration fitness E_{sym}. We utilize Particle Swarm Optimization (PSO) to navigate local minima. The configuration is parameterized as $\boldsymbol{\theta} = [\theta_r, \theta_p, \theta_y, \theta_g]^T$, bounded within physically valid ranges $\Omega = [0, \pi/2] \times [-\pi/2, \pi/2] \times [-\pi/2, \pi/2] \times [0, \pi/2]$.

Each particle i maintains a position $\mathbf{x}_i^t \in \Omega$ and velocity $\mathbf{v}_i^t$. The standard PSO update rules are applied:

$$\mathbf{v}_i^{t+1} = w\mathbf{v}_i^t + c_1 r_1(\mathbf{p}_i - \mathbf{x}_i^t) + c_2 r_2(\mathbf{g} - \mathbf{x}_i^t), \quad (1)$$

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1}, \quad (2)$$



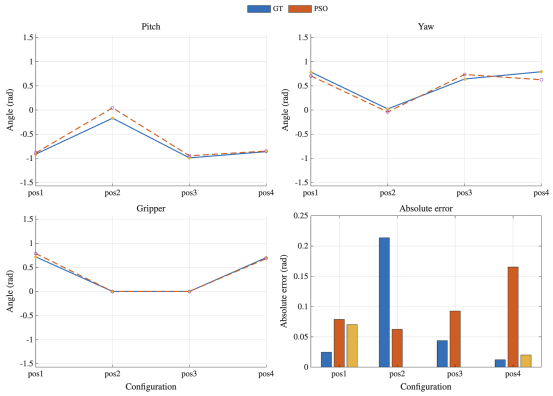
(a) first PSO iteration (b) second PSO iteration (c) third PSO iteration (d) fourth PSO iteration

Fig. 1. Visual representation of the Particle Swarm Optimization (PSO) process. The sequence from (a) to (d) illustrates the progressive refinement of the estimated EndoWrist joint configuration across consecutive optimization steps.

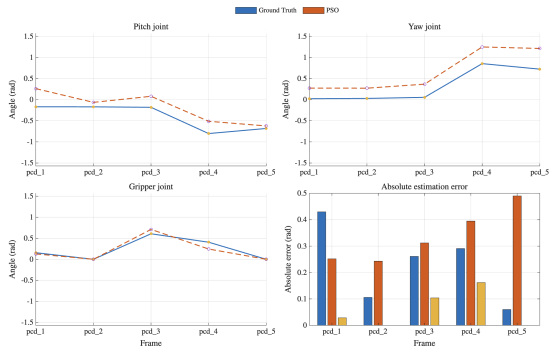
where $\mathbf{p}_i$ is the personal best solution of particle i, $\mathbf{g}$ is the global best solution, and $r_1, r_2 \sim \mathcal{U}(0,1)$. After each update, particle positions are clamped to the feasible domain Ω to enforce joint limits. In our implementation, we use a swarm size of $S = 50$. The PSO parameters are set to $w = 0.7298, c_1 = 1.49445$, and $c_2 = 1.49445$, following the constriction coefficient formulation in [2]. A swarm of 50 particles explores Ω, updating velocities and positions based on personal and global best configurations. Optimization is performed independently per frame and terminates when the global best fitness does not improve over 5 consecutive iterations.

RESULTS AND DISCUSSION

The proposed approach was evaluated on the da Vinci Research Kit (dVRK), leveraging its ground-truth kinematic states to quantitatively assess the estimated joint configurations. Optimization was performed independently per frame using PSO ($S = 50$ particles, $8 \rightarrow 12$ iterations) within the feasible joint domain. To evaluate robustness under varying point cloud quality, we compared two scenarios: (i) baseline point clouds acquired via an Intel RealSense D405 stereo camera, and (ii) point clouds reconstructed from monocular RGB images using Depth Anything v3. In the **RealSense evaluation**, despite high-frequency noise and partial occlusions caused by specular reflections, the point clouds preserved sufficient structural consistency, yielding highly accurate joint estimations across the tested trajectories. Across four evaluated sequences, the mean joint estimation errors were 0.073 rad for the pitch joint, 0.099 rad for the yaw joint, and 0.036 rad for the gripper (see Fig. 2). In the **monocular depth evaluation**, performance was impacted by scale ambiguity and structural distortions inherent to monocular networks. Across five evaluated sequences, the mean joint estimation errors were 0.229 rad for the pitch joint, 0.338 rad for the yaw joint, and 0.059 rad for the gripper (see Fig. 2). The increased error compared to the stereo baseline is primarily attributed to non-metric distortions in the reconstructed point clouds, which affect the ICP-based alignment objective. In conclusion, this work demonstrates that EndoWrist kinematics can be inferred solely from monocular video by formulating pose estimation as a geometry-driven optimization problem. While



(a) RealSense evaluation



(b) Monocular depth evaluation

Fig. 2. The optimization results. a) realSense evaluation: estimated joint value (in orange) compared to dVRK ground truth (in blue), b) monocular depth evaluation: estimated joint values (in orange) compared to dVRK ground truth (in blue).

a performance gap exists between stereo and monocular setups due to the geometric fidelity of the 3D reconstruction, the framework successfully recovers physically plausible joint configurations without direct access to robot encoders.

REFERENCES

[1] L. Yang *et al.*, "Depth anything v3," *arXiv preprint arXiv:2409.09441*, 2024.

[2] M. Clerc and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

Optimizing Path Planning for Enhanced Robotic Ultrasound 3D Reconstruction

Chiara Bernardini[1,2], Kaat Van Assche[1], Ayoob Davoodi[1], Marcello Bonfè[2],
Emmanuel Vander Poorten[1], Gianni Borghesan[1]
[1]Robot-Assisted Surgery (RAS) Group, Department of Mechanical Engineering, KU Leuven, Leuven, Belgium
[2]Department of Engineering, University of Ferrara, Ferrara, Italy

INTRODUCTION

Robotic ultrasound (rUS) is a radiation-free imaging modality that enables real-time, non-invasive visualization of anatomical structures. By improving probe stability and reducing operator dependency, it allows the integration of tracked 2D images into 3D reconstructions, supporting surgical planning and navigation.

However, accurate 3D US reconstruction of complex anatomies remains a challenge due to the strong dependence on the probe–bone incidence angle, with optimal imaging near perpendicular incidence. Therefore, scanning paths with a fixed probe angle often yield incomplete coverage. Van Assche et al. proposed an adaptive replanning framework that improves surface coverage by aligning the probe with the underlying anatomy, based on US images gathered during scanning. However, a key limitation of this approach is that the resulting trajectories often lead to redundant scanning motions and increased acquisition time [1].

In this work we propose a CT-informed path planning framework for rUS that integrates anatomical prior knowledge and explores the use of search and optimisation algorithms commonly employed in robotics for path planning problems [2], [3]. The method generates probe poses that maximize the expected US reflections while minimizing the scanning time, aiming to improve the efficiency and the quality of 3D US reconstruction.

METHOD

The path planning method proposed in this work is a computational framework designed to optimize the scanning path of an ultrasound probe across a single lumbar vertebra.

I. Geometric modeling and bone sampling

A 2D vertebral profile is constructed by projecting the 3D CT point cloud onto the axial plane of the vertebra [1]. The most posterior points are connected using cubic spline interpolation, ensuring continuity of first and second derivatives required for curvature estimation. To identify sampling points while favoring geometrically complex regions, the contour is densely discretized and parametrized by curvature-weighted arc-length s, where local curvature κ (from second-order spline derivatives) and the weighting factor β_{curv} define the sampling density $\rho(s)$:

$$\rho(s) = 1 + \beta_{curv} \frac{\kappa(s)}{\kappa_{max}}$$

Each selected point on the bone cubic spline is defined as $p_b := (x_b, y_b)$. The skin surface is analytically modeled as a wide parabola centered at the spinous process.

II. Pose selection and cost function

For each p_b, the outward normal vector $\vec{n_b}$ is computed to define the optimal probe incidence and generate candidate probe poses on the skin surface. Each pose $M_n := (x_{sn}, y_{sn}, \alpha_{sn})$ is evaluated based on both geometric feasibility and imaging constraints. In particular, a visibility filter (self-occlusion) discards configurations in which the ray between the probe and the target bone point intersects the bone profile at multiple locations, as well as poses lying outside the valid skin domain. The expected image quality is then assessed based on the incidence angle θ between the probe beam λ and $\vec{n_b}$, and each pose is classified into three categories: high ($\theta \leq 8°$), medium ($8° < \theta \leq 15°$), and low ($\theta > 15°$). The method is summarized in Fig. 1.

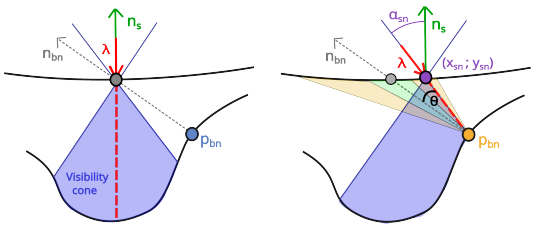


Fig. 1: Generation and evaluation of candidate probe poses. The outward normal at each bone point defines a cone divided in visilibity zones.

A pairwise cost matrix is then constructed to quantify the transition cost between probe poses. The cost function was defined as the time required to move between i and j, considering both translational and rotational motion under bounded velocities. Specifically, the cost function is defined as:

$$C_{ij} = \max\left(\frac{\|p_{sj} - p_{si}\|}{v_{max}}, \frac{|\alpha_{sj} - \alpha_{si}|}{\omega_{max}}\right)$$

in which ω_{max} is derived from the linear velocity limit $v_{\max}$ assuming a probe rotation radius r, i.e., $\omega_{\max} = \frac{v_{\max}}{r}$.

III. Trajectory

The optimal acquisition sequence is formulated as a Traveling Salesman Problem (TSP) over the set of valid poses. The path optimization follows a two-step heuristic. First, an initial tour is generated using a greedy strategy, starting from the leftmost pose (minimum x-coordinate) to

ensure deterministic initialization. The tour is then refined using a simplified Lin–Kernighan-inspired heuristic (LKH-like), which iteratively improves the solution through edge-swapping operations, with the objective of minimizing the total transition cost. The resulting sequence defines the optimized scanning trajectory. During scanning, motion is generated as per [1].

RESULTS

The proposed framework was evaluated using two different sampling densities, generating 20, 50, 70 and 100 candidate n_{imp}. The results are shown in Table I.

TABLE I: Optimization results for different sampling densities.

	Points			
Metric	**20**	**50**	**70**	**100**
Valid poses	20 / 20	48 / 50	68 / 70	97 / 100
High IQ	16	34	45	64
Medium IQ	1	5	8	11
Low IQ	3	9	15	22
Min cost [s]	1.3	0.2	0.3	0.1
Max cost [s]	48.8	48.9	48.8	48.8
Avg cost [s]	24.6	22.8	22.6	22.7
Greedy time [s]	85.5	165.9	211.9	242.0
LKH-like time [s]	122.2	151.4	190.2	214.5

In all the configurations, the majority of valid poses achieved high expected image quality. The corresponding transition cost matrices showed a wide range of motion costs. For the 20-point configuration, the cost ranged from 1.34 s to 48.77 s (average 24.60 s), while for the 100-point case it ranged from 0.1 s to 48.80 s (average 22.70 s). The optimization method was then compared using both a greedy heuristic and the LKH-like approach. For the 20-point sampling, the greedy solution resulted in a total traversal time of 85.51 s, while the refined solution yielded 122.22 s, reflecting the limitations of the simplified heuristic in small-scale problems. In contrast, for the 100-point sampling, the proposed refinement improved the total acquisition time from 242.0 s to 214.5 s. For the sake of clarity, graphical results are shown only for the 20-point configuration (Fig. 2-4).

CONCLUSIONS AND DISCUSSION

This work presents a CT-informed path planning framework for robotic ultrasound imaging of the spine, integrating anatomical prior knowledge with trajectory optimization. The method generates feasible probe poses that maximize the expected image quality while explicitly accounting for motion constraints. By formulating the sequencing problem as a TSP and adopting a Lin–Kernighan-inspired heuristic, the approach enables the computation of efficient scanning trajectories.

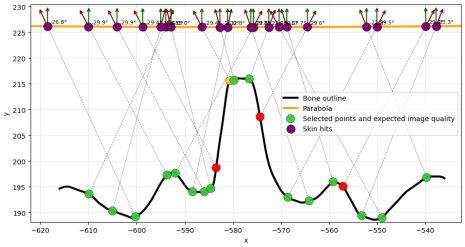


Fig. 2: Candidate skin poses marked with purple dots. Green, yellow, and red circles indicate the bone points reached with high, medium, and low expected image quality respectively.

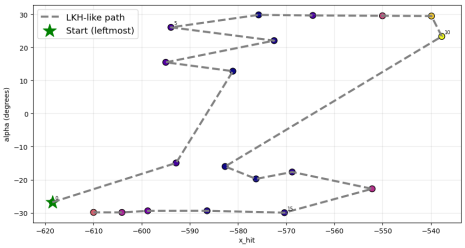


Fig. 3: Sequence of visited poses under the LKH-like heuristic, plotted in pose space (skin x-coordinate vs. probe orientation α). The star marks the starting pose.

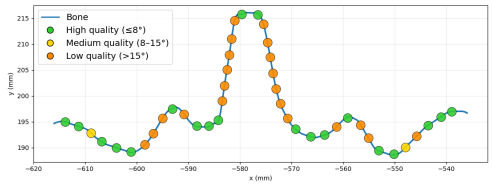


Fig. 4: Bone points observed along the interpolated probe trajectory, colored by expected image quality. Low-quality points cluster where bone curvature is high, limiting achievable incidence angle.

The results indicate that the benefit of the LKH-like optimization becomes more evident as the number of sampling points increases, leading to improved trajectory efficiency: the optimized paths ensure consistent coverage of the bone surface while reducing redundant probe motion, highlighting the potential of the proposed approach for efficient rUS scanning in dense sampling scenarios. Experiments on dedicated phantoms will follow to further validate the framework and assess its performance against reconstruction quality, coverage, and precision metrics [1]. Future work will focus on extending the framework toward an online and adaptive formulation, where the scanning trajectory can be updated in real time based on incoming data, and extending to multiple vertebrae, reporting mean and variability of the outcomes.

REFERENCES

[1] Kaat Van Assche et al. Robotic path re-planning for us reconstruction of the spine. *IEEE Transactions on Medical Robotics and Bionics*, 7(2):755–767, 2025.

[2] K. P. Jayalakshmi et al. A comprehensive survey on coverage path planning for mobile robots in dynamic environments. *IEEE Access*, 13:60158–60185, 2025.

[3] Entesar Alanzi and Mohamed El Bachir Menai. Solving the traveling salesman problem with machine learning: a review of recent advances and challenges. *Artificial Intelligence Review*, 58(9):267, 2025.

Autonomous Endoscopic Navigation Based on Real-Time Lumen Segmentation

Benjamin Calmé and Pietro Valdastri

STORM Lab, Institute of Autonomous Systems and Sensing, University of Leeds, Leeds, UK

INTRODUCTION

Magnetic flexible endoscopy (MFE) decouples intraluminal locomotion from external actuation by steering an internal permanent magnet with a robot-mounted external permanent magnet [1], [2], [3]. This reduces the need for shaft pushing and provides a suitable actuation architecture for supervised autonomy in colonoscopy. A central requirement for autonomous MFE navigation is robust estimation of the forward colon direction from the endoscopic image. Prior work has used threshold-based extraction of the darkest image region as a proxy for the distal lumen, but this strategy is sensitive to folds, specular reflections, bubbles, illumination changes, and sharp turns [1]. When the lumen centre is lost, the robot must stop or wait for correction, increasing the total navigation time. Beyond threshold heuristics, learned lumen following has followed two directions. Depth-based schemes recover a monocular depth map and steer toward its deepest region, extracting a luminal pathway from the reconstructed 3D structure [4]; this aids force-aware manoeuvring but adds a reconstruction and path-extraction stage and runs the control loop at a lower rate. Detection-based schemes track the lumen with bounding boxes and target the box centre [5]; however, a box centre is a coarse proxy that becomes biased for non-convex or partially occluded openings at oblique and hairpin turns. Our alignment law requires only a 2D image-plane target, so we adopt full segmentation: it needs no metric-depth recovery, keeps the loop lightweight, and yields a centroid that follows the visible lumen precisely in the conditions where thresholding and box centres fail. This extended abstract focuses only on autonomous navigation driven by real-time lumen segmentation. We combine a magnetic manipulation model, an image-based alignment controller, and a DeepLabv3+ lumen segmenter trained on phantom and in vivo colonoscopy frames. The objective is to keep the camera aligned with the segmented lumen centre and to advance only when alignment is sufficient, preserving clinician oversight while offloading repetitive centring and progression.

MATERIALS AND METHODS

The robotic platform comprises an MFE tip containing a 10×10 mm internal permanent magnet (IPM), camera, illumination, and service channels, attached to a highly flexible tether. A KUKA LBR Med R820 manipulates a 100×100 mm external permanent magnet (EPM) to generate magnetic forces and torques on the internal magnet (Fig. 1). Hall-effect sensing and an inertial measurement

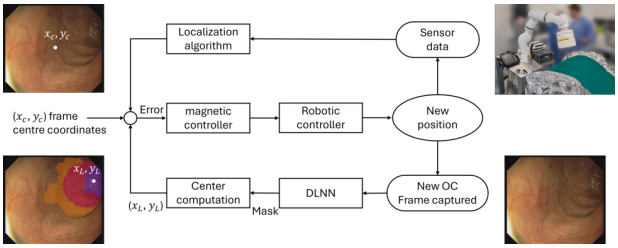


Fig. 1: Control scheme for autonomous navigation based on lumen segmentation and lumen-centre identification

unit provide real-time magnetic localization of the endoscope tip, which closes the control loop together with visual lumen tracking.

Magnetic manipulation model. The magnetic interaction is locally linearized around the current EPM–IPM configuration. For the EPM position $\mathbf{p}_E$ and orientation $\hat{\mathbf{m}}_E$, and the IPM position $\mathbf{p}_I$ and orientation $\hat{\mathbf{m}}_I$, the variation in magnetic force and torque is written as

$$\begin{bmatrix} \delta \mathbf{f}_m \\ \delta \boldsymbol{\tau}_m \end{bmatrix} = \mathbf{J}(\mathbf{p}_E, \mathbf{p}_I, \hat{\mathbf{m}}_E, \hat{\mathbf{m}}_I) \begin{bmatrix} \delta \mathbf{p}_E \\ \delta \hat{\mathbf{m}}_E \end{bmatrix}, \tag{1}$$

where the Jacobian $\mathbf{J}$ is recomputed at each control step. The EPM command is obtained from the pseudoinverse of this local model, combining an orientation-control term with a forward-motion command. In compact form,

$$\begin{bmatrix} \delta \mathbf{p_E} \\ \delta \hat{\mathbf{p}_E} \end{bmatrix} = \mathbf{J}^{\dagger} \begin{bmatrix} 0 \\ \overline{\mathbf{R}_{\mathbf{G}}^{\mathbf{I}} \mathrm{pd}(\delta \mathbf{O_{a,I}}, \overline{\omega_{\mathbf{a,I}}})} \end{bmatrix} + \mathcal{F}(\overline{\delta \mathbf{X_{a,I}}})$$
$$proj_{xy}(\hat{m}_l)(\overline{\delta \mathbf{X_{a,I}}} - \alpha[1 - \frac{F_{mz}}{F_{mz,max}}]) \tag{2}$$

where $\mathbf{R}_G^I$ maps the orientation command to the IPM frame, $\mathrm{pd}(\cdot)$ damps angular error, and the feedforward term $\mathcal{F}$ activates forward progression. If the linearized model approaches a decoupled configuration, a recoupling controller restores EPM–IPM alignment:

$$\delta \mathbf{p}_E = \mathrm{pi} \begin{bmatrix} \mathbf{p_{I,x}} - \mathbf{p_{E,x}} \\ \mathbf{p_{I,y}} - \mathbf{p_{E,y}} \\ \mathbf{h_{reset}} - \mathbf{p_{E,z}} \end{bmatrix} \tag{3}$$

$$\delta \hat{\mathbf{m}}_E = \mathrm{pi}\left(\mathbf{proj}_{xy}(\hat{\mathbf{m}}_I - \hat{\mathbf{m}}_E)\right). \tag{4}$$

Lumen segmentation and navigation control. The perception module uses DeepLabv3+ with a MobileNet encoder to segment the lumen from the endoscopic image. The training set contained 4,891 frames collected from synthetic colon phantoms and in vivo videos, split into 80% training, 10% validation, and 10% testing. Data augmentation included geometric jitter, blur, brightness/contrast

variation, synthetic bubbles, and specularity-like artefacts. The network was trained for 300 epochs using Adam optimization and Dice loss.

For each frame, the binary lumen mask M_t is converted into a centroid $\mathbf{c}_t = [x_l, y_l]^T$. This is compared with the image centre $\mathbf{c}_{img} = [x_c, y_c]^T$ to define the alignment error (Fig. 1)

$$\mathbf{e}_t = \mathbf{c}_{img} - \mathbf{c}_t. \tag{5}$$

The autonomous orientation command sent to the magnetic manipulation controller is

$$\delta\boldsymbol{\theta}_{a,I} = \beta \begin{bmatrix} x_c - x_l \\ y_c - y_l \end{bmatrix}, \tag{6}$$

where β maps pixel error to angular correction. Forward progression is generated from the same image-centering error. The translational input is defined as

$$\delta X_{a,I} = 1 - e^{\kappa\sqrt{(x_c - x_l)^2 + (y_c - y_l)^2}}, \tag{7}$$

where κ tunes the sensitivity of the forward command to the visual alignment error. A temporal filter smooths the centroid before control. The clinician can override the autonomous command, so autonomy acts as a supervised low-level navigation assistant rather than replacing the operator.

RESULTS

On the held-out test set, the lumen segmenter achieved 92.51% mean intersection-over-union and 99.04% pixel accuracy. During replay and robotic navigation, the learned mask provided more stable centroids than the previous threshold-based lumen detector in scenes with folds, illumination changes, partial occlusion, and high-curvature turns. Autonomous navigation was evaluated by comparing the proposed segmentation-based controller with the previous threshold-based approach. Each method was evaluated over 12 trials. The proposed method reduced the mean total navigation time from 498.56 s to 277.98 s and the mean effective navigation time from 408.10 s to 239.29 s. The fully autonomous navigation index (FANI) increased from 63.75% to 78.36%, while the lumen detection index (LDI) increased from 47.31% to 81.79%.

$$\text{FANI} = \frac{\text{fully autonomous navigation time [s]}}{\text{total navigation time [s]}} \tag{8}$$

$$\text{LDI} = \frac{\text{lumen detection time [s]}}{\text{total navigation time [s]}} \tag{9}$$

Additional validation was performed in controlled turning scenarios, where each turn angle was tested over 10 repetitions (Fig. 2). The results confirm the consistent performance improvement achieved by the proposed approach (Tab. I).

These results show that the main benefit of learned lumen segmentation is not only pixel-level accuracy, but also continuity of the control signal. By maintaining a valid lumen centroid through challenging visual conditions, the controller spends less time stopped or misaligned and more time generating safe forward progression.

Alg.	Angle	FANI	LDI
[1]	0°	44.74%	68.63%
[1]	45°	51.22%	74.40%
[1]	90°	64.77%	68.21%
Proposed	0°	65.87%	81.50%
Proposed	45°	79.48%	87.30%
Proposed	90°	99.00%	97.00%

TABLE I: Quantitative evaluation of autonomous navigation performance across turn angles, averaged over 10 trials per angle

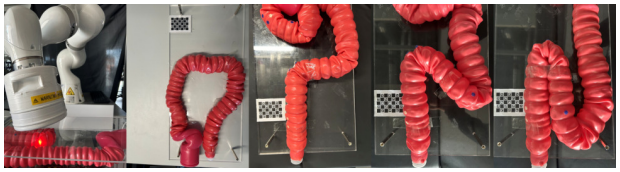


Fig. 2: From left to right: Robotic platform used for the experiments. Then in centre-left, the layout used for the navigation trials, followed by the turn angular progression of the colon phantom starting at 90°, decreasing to 45°, and ending at 0° used during the second set of experiments.

CONCLUSIONS AND DISCUSSION

A DeepLabv3+-based lumen segmenter was integrated with the magnetic manipulation model of an MFE to enable supervised autonomous navigation. Compared with threshold-based lumen extraction, the proposed perception-control loop substantially reduced navigation time and improved both FANI and LDI in the full navigation scenario. The retained manipulation model keeps the approach compatible with existing MFE control infrastructure, while the visual module improves the robustness of the reference direction used for low-level alignment.

Remaining limitations include the need for more heterogeneous in vivo training data, explicit treatment of localization uncertainty, and improved fallback behaviour when masks become unreliable. Future work will focus on expanding the lumen dataset, adding confidence-aware control arbitration, and validating the approach across more anatomically variable phantom and preclinical scenarios.

REFERENCES

[1] J. W. Martin, B. Scaglioni, J. C. Norton, V. Subramanian, A. Arezzo, K. L. Obstein, and P. Valdastri, "Enabling the future of colonoscopy with intelligent and autonomous magnetic manipulation," *Nature Machine Intelligence*, vol. 2, pp. 595–606, 2020.

[2] N. J. Greenidge, B. Calmé, A. C. Moldovan, B. Abaravicius, J. W. Martin, N. Marahrens, J. Woolfrey, B. Scaglioni, D. S. Chathuranga, S. Mitra, S. Cochran, and P. Valdastri, "Harnessing the oloid shape in magnetically driven robots to enable high-resolution ultrasound imaging," *Science Robotics*, vol. 10, no. 100, p. eadq4198, 2025.

[3] Y. Li, Y. Sun, W. Li, R. Xu, W. Y. Ng, P. W. Y. Chiu, and Z. Li, "Design and vision-based control of a compact magnetically actuated soft-tethered colonoscope for automatic manipulation with safety constraints," *IEEE/ASME Transactions on Mechatronics*, 2026.

[4] T. Yang, Y. Yang, P. Wang, Y. Cao, Z. Yang, and H. Liu, "A lumen-adapted navigation scheme with spatial awareness from monocular vision for autonomous robotic endoscopy," *Robotics and Autonomous Systems*, 2023.

[5] B. Al-Bander, A. Mathew, L. Magerand, E. Trucco, and L. Manfredi, "Real-time lumen detection for autonomous colonoscopy," in *Imaging Systems for GI Endoscopy and Graphs in Biomedical Image Analysis (ISGIE 2022 / GRAIL 2022)*. Springer, 2022.

Remote Center of Motion Evaluation of a Compliant Joint for Neuroendoscopy

Federico Mariano[1,2], Elena De Momi[2], Giovanni Berselli[1,3], and Leonardo S. Mattos[1]

[1] *Biomedical Robotics Lab, Advanced Robotics, Istituto Italiano di Tecnologia, Genoa, Italy*

[2] *Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy*

[3] *DIME, Università di Genova, Genoa, Italy*

INTRODUCTION

Neuroendoscopy represents a specialized minimally invasive surgical approach that plays an important role in the treatment of hydrocephalus and selected intraventricular, skull base, and other midline intracranial lesions. Compared with conventional open approaches, neuroendoscopy offers several potential advantages, including a smaller surgical corridor, reduced brain retraction, limited tissue manipulation, improved illumination and visualization of deep-seated structures. These features may result in lower surgical morbidity and, in some indications, shorter operative time and hospital stay [1].

To safely support neuroendoscopic procedures, robotic systems must allow the instrument to pivot about a Remote Center of Motion (RCM) located at the burr-hole, thereby allowing accurate endoscope orientation with respect to selected insertion trajectory. Several RCM solutions have been proposed in the literature, including parallelogram-based mechanisms [2], spherical wrists [3], and software-enforced constraints [4]. However, these approaches may suffer from drawbacks, such as bulky architectures, backlash, friction, and sensitivity to calibration and assembly errors.

In this context, compliant RCM mechanisms are particularly attractive because their monolithic structure eliminates joint clearances, reduces friction, and enables compact designs with smoother motion and intrinsic kinematic guidance [5]. Among these, Rommers et al. introduced the Tetra II joint, a compliant mechanism in which the RCM is maintained fixed in space through a tetrahedral architecture [6].

This work investigates the use of compliant mechanisms in neuroendoscopy. Specifically, it studies the performance of a novel off-axis compliant RCM joint with respect to neuroendoscopy requirements. The mechanism was designed based on a two-stage analytical and numerical procedure aimed at achieving near-isotropic stiffness while minimizing parasitic displacement at the RCM [7]. Experiments were then carried out to evaluate the accuracy of the mechanism, with particular focus on the parasitic displacement at the RCM, which must stay within 1 mm for its safe use in neuroendoscopy.

MATERIALS AND METHODS

A dedicated experimental setup was developed to evaluate the parasitic displacement of the RCM of the proposed compliant joint. The mechanism was actuated through two orthogonal ball screws, enabling controlled and repeatable motion during testing. The joint was analyzed by varying of 30° the angular coordinate θ in the x–z plane, consistently with the kinematic description reported in Fig. 1. To measure the motion of the endoscopic shaft, a mock endoscope, shown in Fig. 2, was 3D printed and instrumented with two NDI Aurora electromagnetic sensors positioned at its proximal and distal ends. The use of two sensors was motivated by the limited accuracy of the single sensor localization with respect to the small parasitic displacements expected at the RCM. Their combined measurements improve the robustness of the measurement and reduce the influence of sensor noise on RCM estimation.

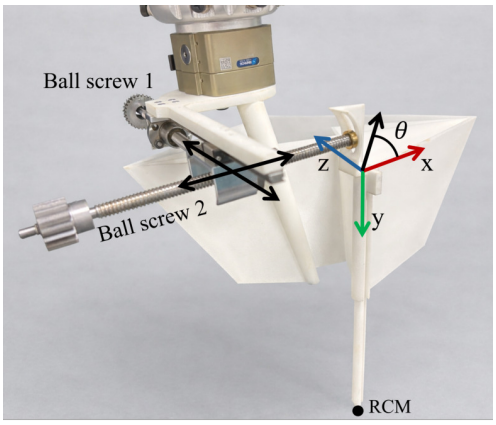


Fig. 1: Off-axis compliant joint integrated with two orthogonal ball screw.

For each test, the endoscope axis was reconstructed using the two electromagnetic sensors. The positions before and after multiple angular displacement of approximately 4.5° were obtained. This angular displacement was selected as representative of the motion required to reach the relevant workspace in third-ventricle neuroendoscopy [8]. The acquired data were processed in MATLAB to identify the pair of closest points between two reconstructed axes, corresponding to the estimated RCM location. The distance between these points was then taken as the parasitic RCM error. For each angular deviation, the mean error was computed based on eight measurements. This was repeated for 12 different angular configurations of the joint, with θ varying from 0° to 330° in 30° steps.

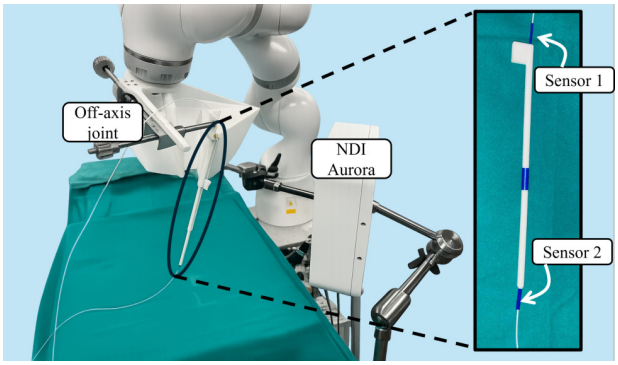


Fig. 2: Setup used for finding the displacement of the RCM error.

TABLE I: Parasitic RCM error measured at different values of θ.

θ [deg]	Mean ± SD [mm]	Min [mm]	Max [mm]
0	0.0597 ± 0.0269	0.0098	0.0812
30	0.1516 ± 0.1051	0.0193	0.3161
60	0.1129 ± 0.0755	0.0411	0.2647
90	0.2198 ± 0.1532	0.0346	0.4016
120	0.1566 ± 0.0767	0.0660	0.2556
150	0.1165 ± 0.0887	0.0059	0.2809
180	0.0690 ± 0.0297	0.0411	0.1188
210	0.1781 ± 0.0605	0.0814	0.2849
240	0.1672 ± 0.1135	0.0027	0.3077
270	0.1218 ± 0.1132	0.0054	0.3185
300	0.0760 ± 0.0353	0.0351	0.1346
330	0.1007 ± 0.0438	0.0409	0.1699

RESULTS

The parasitic displacement results are summarized in Table I and Fig. 3. They show sub-millimetric displacement of the RCM for all the values of θ, with an overall mean error of 0.1275 mm and a maximum measured value of 0.4016 mm.

At the same time, the measurements highlight a non-uniform behavior across the explored angular configurations. The lowest mean parasitic displacement values were measured at 0° and 180°, corresponding to 0.0597 mm and 0.0690 mm, respectively, while the highest mean error was observed at 90°, with a value of 0.2198 mm. Relatively larger mean errors were also found at 210° and 240°, equal to 0.1781 mm and 0.1672 mm, respectively.

DISCUSSION AND CONCLUSION

This work presented the experimental evaluation of an off-axis compliant RCM joint designed for a future application in robot-assisted neuroendoscopy. The proposed experimental setup, based on two orthogonal ball screws and electromagnetic tracking of a mock endoscope, allowed the parasitic displacement at the RCM to be quantified over the investigated workspace in a controlled and repeatable manner.

The experimental results showed RCM displacement values well within the threshold limit of 1 mm. At the same time, the non-uniform behavior observed across the explored angular configurations suggests that the accuracy of the mechanism is affected by its angular configuration,

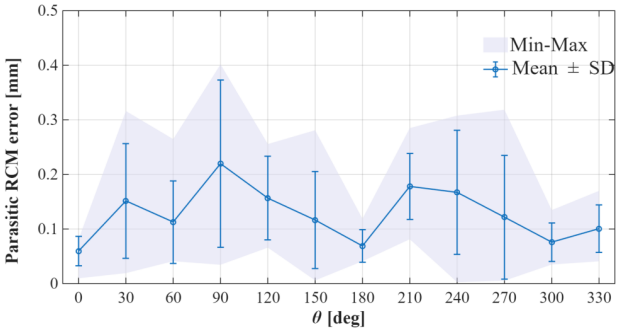


Fig. 3: Parasitic RCM error across angular configurations.

possibly due to residual asymmetries in the compliant architecture. Nevertheless, the measured parasitic displacement remained limited, confirming the capability of the joint to preserve the RCM with good accuracy.

From an application perspective, these results are encouraging for the intended application. In fact, the possibility of combining compactness, reduced friction, and the absence of backlash with a limited parasitic error represents a significant advantage over more conventional rigid-joint solutions. In this sense, the present study supports the potential of compliant architectures as a promising option for robot-assisted neurosurgical procedures.

Future work may include motorizing the actuation of the device. In addition, pre-clinical tests could be carried out on a custom-made phantom head to validate the proposed mechanism in a more realistic surgical scenario and further assess its suitability for neuroendoscopic applications [9].

REFERENCES

[1] F. Yang, W. Xu, X. Tang, Y. Yang, B. A. Ku, Y. Zhang, X. Yang, W. Xie, and X. Hui, "The efficacy of neuroendoscopic surgery treating patients with thalamic hemorrhage accompanied by intraventricular hematoma," *Frontiers in Surgery*, vol. 11, p. 1472830, 2024.

[2] A. Singh et al., "Sensitivity analysis of a double-parallelogram based rcm mechanism used for mis robots," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 237, no. 20, pp. 4813–4827, 2023.

[3] M. A. Laribi et al., "A design of slave surgical robot based on motion capture," in *2012 IEEE international conference on robotics and biomimetics (ROBIO)*. IEEE, 2012, pp. 600–605.

[4] M. M. Marinho et al., "A programmable remote center-of-motion controller for minimally invasive surgery using the dual quaternion framework," in *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, 2014, pp. 339–344.

[5] P. Pluimers et al., "A compliant on/off connection mechanism for preloading statically balanced compliant mechanisms," vol. 4, 08 2012.

[6] J. Rommers et al., "A new type of spherical flexure joint based on tetrahedron elements," *Precision Engineering*, vol. 71, 03 2021.

[7] F. Mariano, E. D. Momi, G. Berselli, J. Jovanova, J. L. Herder, and L. S. Mattos, "Off-axis compliant rcm joint with near-isotropic stiffness and minimal parasitic error," 2026. [Online]. Available: https://arxiv.org/abs/2603.28240

[8] F. Duffner, H. Schiffbauer, D. Glemser, M. Skalej, and D. Freudenstein, "Anatomy of the cerebral ventricular system for endoscopic neurosurgery: a magnetic resonance study," *Acta neurochirurgica*, vol. 145, no. 5, pp. 359–368, 2003.

[9] F. Mariano, C. Abello, N. Kamal, G. Berselli, E. De Momi, G. Piatelli, and L. S. Mattos, "Design and fabrication of a phantom head for robotic neurosurgery simulation," in *European Robotics Forum*. Springer, 2024, pp. 43–47.

Computationally efficient workspace boundary determination for two-segment continuum robots

Joseph Massin and M. Taha Chikhaoui

Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, F-38000 Grenoble, France

INTRODUCTION

Due to their slender shape and flexibility, continuum robots are widely used in minimally invasive surgery. They are commonly task-specifically designed, which requires foremost optimization of their reachable workspace. Due to the interdependence of the robot parameters, computationally expensive metaheuristic optimization algorithms, such as evolutionary or particle swarm algorithms, are employed. During the optimization process, the workspace is evaluated numerous times, often with brute force [1] or random sampling [2] methods. The overall computational cost is reduced for continuation approaches where only the workspace's boundaries are estimated [3]. This abstract aims to provide a simpler, more efficient method for computing those boundaries, applicable to several architectures of two-segment continuum robots.

MATERIALS AND METHODS

Each robot segment consists of a curved portion and possibly a straight part. They can be fully retracted and deployed with regard to the actuation unit. Segment curvatures can be either pre-defined, e.g. Concentric Tubes Continuum robots (*CTCR*), or controlled by actuators, e.g. Tendon Actuated Continuum Robots or Concentric Push-Pull Robots (*TACR*, resp *CPPR*). Only robots with constant curvatures are considered, and the curvatures are assumed to lie in a plane that can be, due to actuation, rotated around an axis. As the workspace is radially symmetric, its boundaries can be represented as curves in a plane, and the 3D workspace is fully defined. Since no external loads are considered, such robots can be modeled with piecewise constant curvature, as:

$$\begin{cases} \dot{r}_y(s) = -\sin(\varphi_x(s)) \\ \dot{r}_z(s) = \cos(\varphi_x(s)) \end{cases} \quad \text{and} \quad \begin{cases} \dot{\varphi}_x(s) = \kappa_x(s) \\ \dot{\kappa}_x(s) = 0, \end{cases} \tag{1}$$

where $r_y(s)$ and $r_z(s)$ are the robot position in the plane $(\mathbf{y}, \mathbf{z})$, and $\varphi_x(s)$ its orientation around the $\mathbf{x}$ axis, while $\kappa_x(s)$ is the curvature of the portion resulting from the elastic equilibrium of component segments. The dot notation denotes differentiation with respect to the arc length s. To Eq. (1) must be added the initial conditions $r_y(0) = r_z(0) = 0$, and $\varphi_x(0) = 0$, and the shape's continuity between portions. A differential analysis was performed based on Eq. (1), and results are validated against more than 3×10^6 configurations of brute force workspace generated from [4]'s model, considered as the ground-truth.

This work aims to provide optimal actuations that yield the workspace's external boundary, applicable for all of the aforementioned robots. The latter is separated into two parts, hereafter referred to as the *lateral* and *distal* boundaries, which are not defined here geometrically, but rather by actuation sequences. The lateral boundary is obtained by varying the translation actuators while maintaining the segments bent in a common plane. The distal boundary is obtained by adjusting the segments' rotation actuation when the robot is fully deployed. For robots with pre-defined curvature (CTCR), an *inner* boundary of unreachable points can appear within the workspace, but is not addressed here.

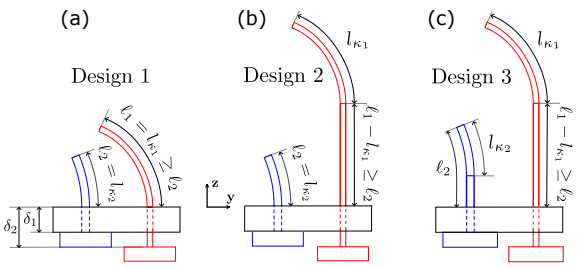


Fig. 1: Parameterization of the three pairs of segments.

RESULTS

Three different pairs of segments, shown in Fig. 1, are studied. The inner segment (segment 1) is longer than the outer one (segment 2). Both are defined by their total length L_i, and the length of their curved portion $l_{\kappa i}$, for $i = 1, 2$. The maximal length of the i^{th} segment is $\ell_i = L_i - \delta_i$, where δ_i is the minimal distance between segment i's carriage and the base plate, see Fig. 1. Small lengths ε_1 and ε_2 prevent the segments from retracting into each other and into the actuation unit, respectively. Thus one has $\varepsilon_2 \leq l_{d_2} \leq \ell_2$ and $l_{d_2} + \varepsilon_1 \leq l_{d_1} \leq \ell_1$, where l_{d_i} is the deployed length of the i^{th} segment.

a) Lateral boundary: To reach this boundary, segments should be fully bent (due to maximum actuation for TACR and CPPR or due to pre-curvature for CTCR) and the curvature of both segments (referred to as κ_i^*) is positive and lies in a same plane. To each robot's deployed length l_{d_1} corresponds a unique point on the lateral boundary, that can be reached by optimally deploying the outer segment alongside the inner, *i.e.* by finding the optimal value of $l_{d_2}(l_{d_1})$.

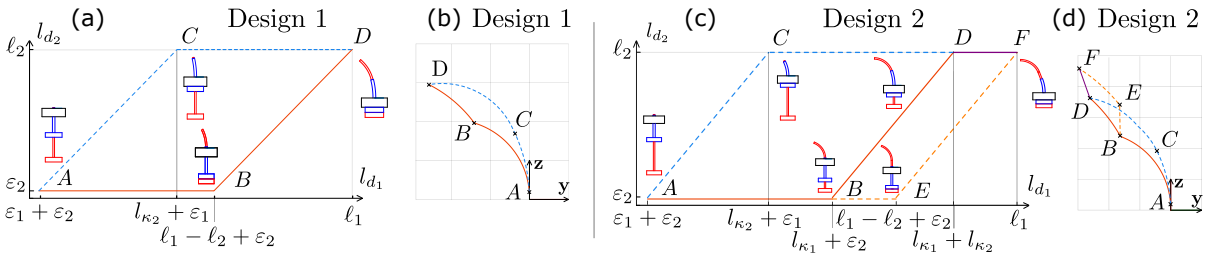


Fig. 2: (a) & (c) Sequence of deployed lengths, (b) & (d) Corresponding end-effector's position, for Designs 1 and 2.

In Design 1, (Fig. 1a), both segments have no straight portion, see [5] for such an example. The optimal value of l_{d_2} only depends on the relative value of segment curvatures. If $\kappa_1^* \leq \kappa_2^*$, both segments should deploy at once until the segment 2 reaches its maximal length. Then, only segment 1 can deploy. If $\kappa_1^* > \kappa_2^*$, the lateral boundary is obtained by deploying at first only segment 1, until its carriage get in contact with segment 2. Then, both segments have to deploy together until they reach their maximal length. This sequence (ABD in Fig. 1a) defines a *path* and is drawn in solid lines. The corresponding lateral boundary is shown in Fig. 2b. Note that Figs. 1, to 4 are illustrated with $\kappa_1^* > \kappa_2^*$.

In Design 2, (Fig. 1b), segment 1 has a straight portion, as in [6]. As previously, if $\kappa_1^* \leq \kappa_2^*$, both segments should be deployed until segment 2 reaches its maximum length, then keep on deploying segment 1. If $\kappa_1^* > \kappa_2^*$, the boundary is at first obtained by deploying only segment 1 until its curved portion is totally exposed. Then, both segments should be simultaneously deployed until the curvature of segment 2 is also fully exposed. Finally, the remaining of segment 1 should be deployed. This yields the path $ABDF$ (Fig. 2c), whose corresponding lateral boundary is drawn in Fig. 2d.

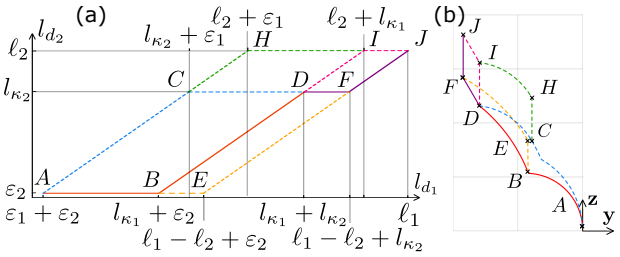


Fig. 3: (a) Sequence of deployed lengths, (b) Corresponding end-effector's positions, for Design 3.

Design 3, (Fig. 1c), consists of two segments with both straight and curved portions, as in [7]. Similarly to Design 2, the lateral boundary is firstly obtained by deploying the curved portion of both segments (therefore following path ABD or ACD in Fig. 3a, depending on the relative value of κ_1^* and κ_2^*). Once point D is reached, the remaining part of the lateral boundary is obtained by following the path DEF, see Fig. 3b, no matter the relative amplitude of the two segments' curvature.

b) Distal boundary: If segments have a pre-defined curvature (e.g. CTCR), the distal boundary is obtained first by rotating segment 1 relatively to segment 2, then by rotating both segments altogether until the robot's tip is brought back to the plane $(\mathbf{y}, \mathbf{z})$.

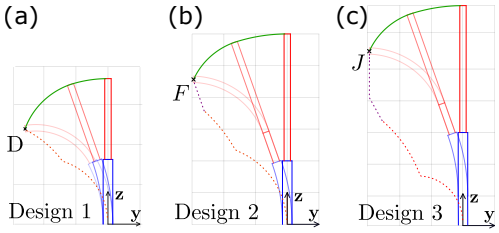


Fig. 4: Distal boundaries of the three designs.

If robots have actuated curvatures (e.g. TACR or CPPR), such boundary is obtained from their fully deployed configuration by simply straightening the segments' curvatures. From Eq. (1), it can be shown that the distal boundary is obtained as a sequence of first straightening up segment 1, then segment 2, as shown in green in Fig. 4 for all designs.

CONCLUSIONS AND DISCUSSION

An efficient strategy is proposed to evaluate the 3D workspace boundaries of continuum robots through the determination of an optimal combination of actuation. The method is of $O(n)$ numerical complexity and applies to several types of robots and to different segment combinations. Although this work focuses on two-segment robots, the results can readily be extended to three-segment ones, at the cost of a high number of sequences.

REFERENCES

[1] A. Ulrich *et al.*, "Design optimization of a concentric tube continuum robot for surface ecog brain implant insertion in a highly constrained environment," *CRAS, Lisbon, Portugal*, 2025.

[2] C. Girerd and T. K. Morimoto, "Design and control of a hand-held concentric tube robot for minimally invasive surgery," *IEEE Trans. Robot. 37(4): 1022-1038*, 2020.

[3] W. Amehri, G. Zheng, and A. Kruszewski, "Workspace boundary estimation for soft manipulators using a continuation approach," *IEEE Robot Autom Let. 6(4): 7169-7176*, 2021.

[4] Q. Boyer *et al.*, "On high performance control of concentric tube continuum robots through parsimonious calibration," *IEEE Robot. Autom. Lett. 9(11): 9621-9628*, 2024.

[5] E. Amanov, J. Granna, and J. Burgner-Kahrs, "Toward improving path following motion: Hybrid continuum robot design," in *IEEE Int. Conf. Robot. Autom*, 2017, pp. 4666–4672.

[6] J. A. Childs and C. Rucker, "A kinetostatic model for concentric push–pull robots," *IEEE Trans. Robot. 40: 554-572*, 2023.

[7] J. Ha, G. Fagogenis, and P. E. Dupont, "Modeling tube clearance and bounding the effect of friction in concentric tube robot kinematics," *IEEE Trans. Robot. 35(2): 353-370*, 2019.

Acknowledgments This work is supported by the research projects funded by grants ANR-24-CE33-4611, ANR-21-ESRE-0015, and ANR-23-IACL-0006.

Deformation-Aware Autonomous Circular Cutting in Robot-Assisted Surgery Using MFT-Based Tracking and Online Scissor Pose Planning

Yuriko Iyama, Pietro Valdastri, and Dominic Jones

School of Electronic and Electrical Engineering, University of Leeds, Leeds, UK

INTRODUCTION

Tissue cutting is a fundamental surgical skill required in nearly all procedures, as reflected in standardized training curricula such as the Laparoscopic Skills Assessment (LapPass), which includes "Cutting an Accurate Disc" as a core module [1]. Despite its importance, autonomous tissue cutting remains underexplored in robot-assisted minimally invasive surgery (RAMIS) compared to tasks such as needle insertion and knot tying [2].

The primary challenge arises from the deformable and viscoelastic properties of soft tissue, which cause the target cutting line to change continuously, making it difficult to maintain accurate correspondence between the planned trajectory and the instantaneous tissue configuration.

Although prior studies have addressed this problem, several limitations remain. Murali *et al.* [3] achieved automated cutting along predefined trajectories but did not account for tissue deformation. Kam *et al.* [4] enabled deformation tracking and occlusion-aware online planning using RGB-D sensing; however, their approach relies on an electrosurgical probe that does not require pose planning. Wang *et al.* [5] demonstrated autonomous dissection using mesh-based tracking and visual-topological feedback, but did not consider online scissor pose planning.

To address these limitations, we propose a deformation-aware autonomous framework for 2D circular cutting, integrating RGB-D sensing, Multi-Flow Dense Tracking (MFT) [6], Thin Plate Spline (TPS) interpolation, and online scissor pose planning. As an initial validation, we evaluate the framework under static conditions to establish baseline accuracy prior to deployment in dynamic tissue environments.

MATERIALS AND METHODS

Fig. 1 presents the proposed framework. RGB-D images are processed using the MFT [6] to track surface points without markers. TPS interpolation then estimates the current cutting line from the spatial correspondence between tracked points and the cutting-line points defined in the initial frame. The estimated points are passed to the scissors pose planning and execution module. The method assumes that closing the scissors at a planned pose reliably cuts up to the corresponding indexed point; since no explicit cut-state feedback exists, an incomplete cutting detection module is incorporated as a final verification step to identify and recover any residual uncut regions.

MFT was selected over RAFT, CoTracker, and Optical Flow for its superior robustness to occlusion and tracking

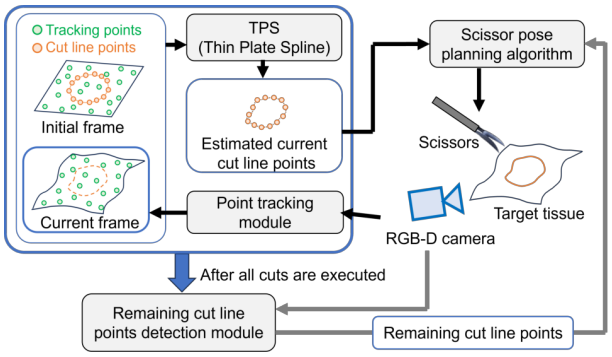


Fig. 1: Overview of proposed framework

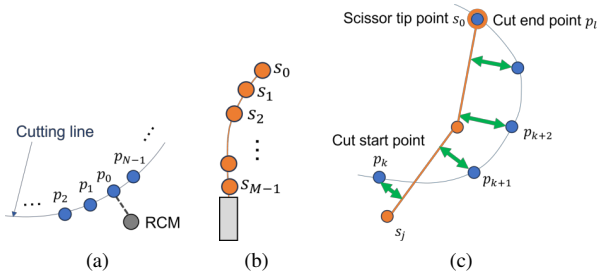


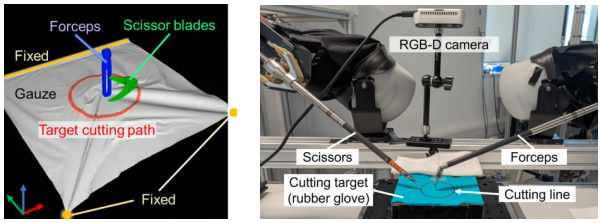
Fig. 2: Indexed point representation of the target cutting line (a) and scissor blade (b) and scissors pose planning algorithm (c)

accuracy. TPS interpolation then reconstructs the full cutting line from the tracked points, naturally accommodating deformation or displacement caused by tissue deformation.

The online scissors pose planning algorithm operates on N indexed points along a circular target line, with the curved blade modeled as a discrete point set (Fig. 2(a)(b)). The scissors are assumed perpendicular to the XY plane, reducing the pose to a tip position and Z-axis rotation. The algorithm minimizes the distance between the blade and the target line, advancing the cutting index until a threshold is exceeded to determine orientation and cutting extent per stroke.

After all planned cuts, color-based segmentation is applied to RGB images to detect residual incisions. Regions where the underlying substrate is exposed are classified as successfully cut, while areas retaining the target color indicate incomplete incision. If such regions are detected, the pose planning algorithm is re-invoked.

Validation was carried out in three stages using two environments (Fig. 3): a physics-based SOFA simulation



(a) Simulation environment (b) Physical robot experimental setup

Fig. 3: Simulation and experimental setups

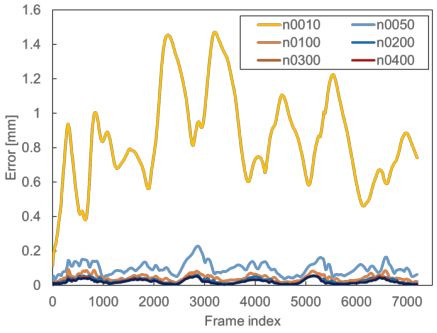
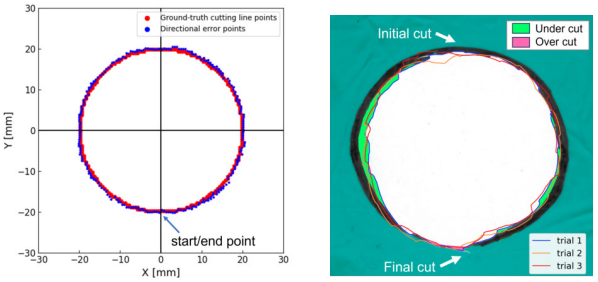


Fig. 4: TPS estimation error vs. number of tracking points



(a) Scissor pose planning error (b) Result on the dVRK

Fig. 5: Experimental results

with FEM, in which mesh incision was realized via the SOFA Carving plugin, and a physical da Vinci Research Kit (dVRK) setup. First, to assess the accuracy of TPS-based cutting line estimation, a cloth substrate was deformed sinusoidally (amplitude: 10 mm, period: 12 s) in simulation, and estimation error was measured as a function of the number of tracking points. Second, to evaluate the scissor pose planning algorithm, a pattern cutting task was performed in simulation, and the nearest-neighbor distance from each point on the target cutting line to the actual cut trajectory was measured as the error metric. Third, to validate the complete framework, the pattern cutting task was executed on the dVRK over three trials using 300 tracking points. To enable direct comparison with human performance, we adopted the composite z-score metric from [7], which jointly captures task completion time and cutting accuracy relative to human operators. Scissor tip trajectory length was additionally recorded and benchmarked against the same human dataset.

RESULTS

TPS estimation accuracy is shown in Fig. 4. With 100+ tracking points, the mean error fell below 0.05 mm (range: 0.021–0.857 mm across 10–500 points), confirming sufficient accuracy relative to the 2 mm target incision width.

The evaluation of the scissor pose planning algorithm (Fig. 5(a)) achieved a mean 0.28 mm (max: 0.86 mm), with all errors below 1 mm. Full framework results are shown in Fig. 5(b). The composite scores (mean ± SD: 0.01 ± 0.39) were comparable to the untrained human mean (0.00) [7]. The task time component (mean: -0.43) was below the human mean, whereas the error zone pixel component (mean: 0.44) remained above it, indicating that task speed was achieved at the cost of cutting accuracy. Scissor tip trajectory length (mean: 23.6 cm) was less than half the human mean (55.6 cm; std: 18.6 cm), indicating geometrically efficient motion.

CONCLUSIONS AND DISCUSSION

This study presented an autonomous framework for deformation-aware 2D circular cutting with an incomplete-cut recovery mechanism. As an initial validation under static conditions, simulation confirmed TPS estimation error below 0.05 mm and scissor pose planning error below 1 mm. Physical dVRK experiments yielded a composite z-score of 0.01 ± 0.39, comparable to the untrained human mean (0.00). While the task-time component (mean z: -0.43) fell below the human mean, the cutting-error component (mean z: +0.44) exceeded it, indicating that speed came at the cost of accuracy. Hand-eye calibration and robot kinematic errors are suggested as potential sources of this reduced cutting accuracy. Scissor tip trajectory length (23.6 cm) was less than half the human mean (55.6 cm), confirming geometric efficiency. Limitations of the current study include the use of phantom rather than biological tissue, restriction to 2D planar cutting, and evaluation under static rather than dynamic deformation conditions. Validation on ex vivo tissue and extension to dynamic environments and 3D cutting represent the immediate next steps.

REFERENCES

[1] The Association of Laparoscopic Surgeons of Great Britain and Ireland, *LapPass Handbook*, The Association of Laparoscopic Surgeons of Great Britain and Ireland, 2023. [Online]. Available: https://www.alsgbi.org/wp-content/uploads/2023/11/ALSGBI-LapPass-Handbook-National.pdf

[2] T. E. Shkurti and M. Cenk Çavuşoğlu, "A Systematic Review of Task Automation in Surgical Robotics," *IEEE Trans. Med. Robot. Bionics*, vol. 7, no. 3, pp. 863–880, 2025.

[3] A. Murali *et al.*, "Learning by observation for surgical subtasks: Multilateral cutting of 3D viscoelastic and 2D Orthotropic Tissue Phantoms," in *2015 IEEE ICRA*, 2015, pp. 1202–1209.

[4] M. Kam *et al.*, "Autonomous Closed-Loop Control for Robotic Soft Tissue Electrosurgery Using RGB-D Image Guidance," *IEEE Trans. Med. Robot. Bionics*, vol. 7, no. 3, pp. 1041–1050, 2025.

[5] C. P. Wang *et al.* (2025) Feedback Matters: Augmenting Autonomous Dissection with Visual and Topological Feedback.

[6] M. Neoral, J. Šerých, and J. Matas, "MFT: Long-Term Tracking of Every Pixel," in *2024 WACV*, 2024, pp. 6823–6833.

[7] Y. Sharon *et al.*, "Augmenting Robot-Assisted Pattern-Cutting With Periodic Perturbations: Can We Make Dry Lab Training More Realistic?" *IEEE Trans. Biomed. Eng.*, vol. 72, no. 1, pp. 264–275.

Delay Compensation for Micro-Macro Bilateral Teleoperation in Robot-Assisted Transoral Laser Microsurgery

Ayoob Davoodi[1], Sukrit Prasarnkleo[1], Mouloud Ourak[1], Gianni Borghesan[1], and Emmanuel Vander Poorten[1]

[1]*Department of Mechanical Engineering, KU Leuven, Leuven, Belgium.*

INTRODUCTION

Transoral Laser Microsurgery (TLM) as treatment for Upper aero-digestive tract (UADT) cancers has become a preferred minimally invasive treatment approach for glottic and supraglottic lesions due to its ability to access the surgical site through the natural oral cavity without external incisions [1]. However, conventional TLM relies on long rigid instruments operated under microscopic visualisation, resulting in limited dexterity, poor ergonomics, and the absence of haptic feedback. To improve surgical manipulation, the CARMA (CollAborative Robot MAnipulator) platform was developed using a KUKA LBR Med14 robot integrated with a microsurgical forceps end-effector and a Remote Centre of Motion (RCM) constraint, see Fig. 1. The surgeon operates the system remotely through a Haptic Desktop 6D interface. Previous validation studies demonstrated accurate path-following performance; however, residual temporal delays of approximately 0.12–0.22 s were observed even in unilateral operation, highlighting the importance of delay compensation in future bilateral implementations. Introducing haptic force feedback in bilateral teleoperation can improve interaction awareness and reduce the risk of excessive tissue forces. However, communication delays may destabilise the system and degrade transparency, particularly in micro–macro teleoperation architectures where motion and force scaling amplify delay-induced errors. Existing delay compensation approaches have not been extensively investigated under combined micro–macro scaling, time-varying delay, and switched tissue interaction conditions.

Therefore, this work proposes an Extended Kalman Filter (EKF)-based delay observer combined with a first-order predictive compensation framework for bilateral robotic microsurgery. A switched Kelvin–Boltzmann/Hunt–Crossley contact model is incorporated to emulate transitions between free-space motion and soft-tissue interaction. The proposed framework is evaluated in multiple simulation scenarios to assess delay-estimation accuracy, tracking performance, and closed-loop response under varying communication delays.

MATERIALS AND METHODS

The proposed bilateral control framework, shown in (Fig. 1, Bottom), is developed for future implementation on the physical CARMA platform (Fig. 1, top) and is evaluated in simulation in the present study. Position commands

This project received funding from European Union through the AIRCARE innovation action project (no. 101137426).

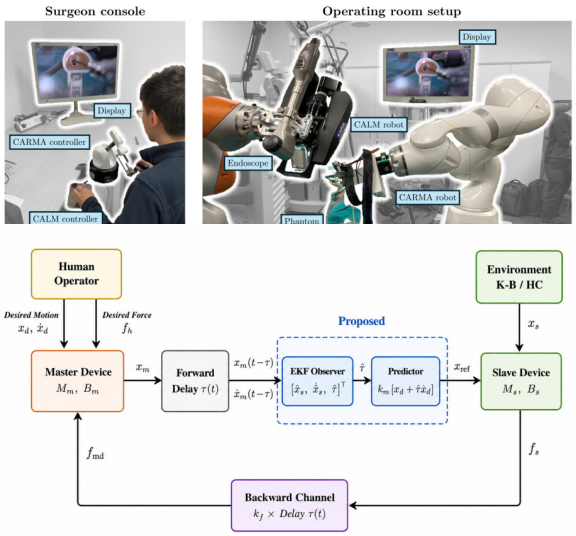


Fig. 1: CARMA bilateral teleoperation system.

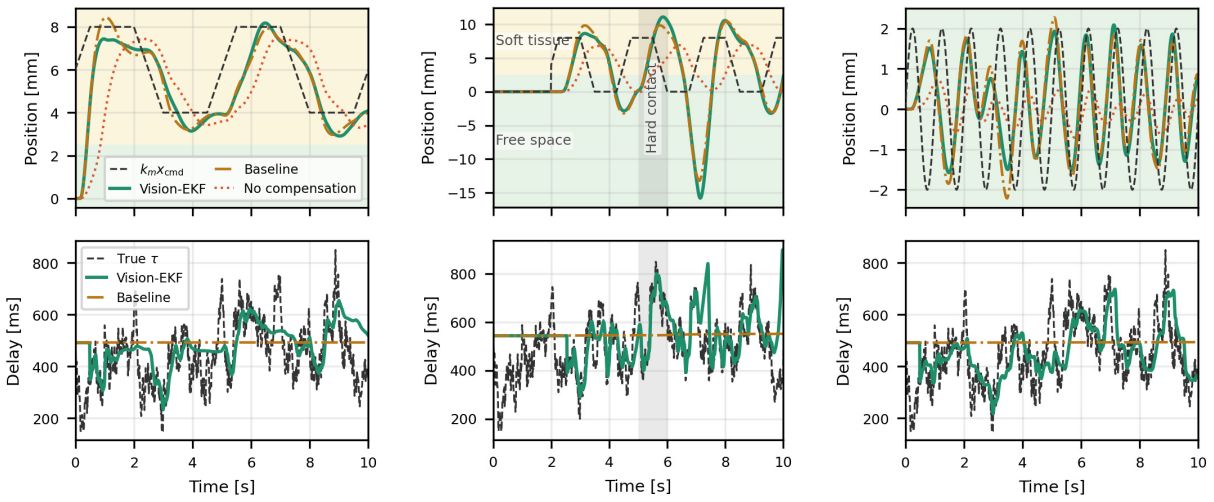
generated at the master side are scaled and transmitted to the slave robot through a delayed communication channel, while interaction forces measured at the slave side are reflected back to the operator through a haptic feedback loop. Motion downscaling and force upscaling are applied to achieve precise microsurgical manipulation and enhanced force perception. In the investigated configuration, motion downscaling ($k_m = 0.2$) and force upscaling ($k_f = 10$) are independently selected; since these gains are not power-preserving, the resulting closed-loop response is evaluated numerically rather than claiming passivity.

1) System Dynamics: The master and slave subsystems are modelled as Cartesian impedance systems in tool-tip space:

$$\begin{aligned} M_m\ddot{x}_m + B_m\dot{x}_m &= f_h(t) - k_f f_s(t-\tau), \\ M_s\ddot{x}_s + B_s\dot{x}_s &= u_s - f_s, \end{aligned} \tag{1}$$

where M_m and B_m denote the master inertia and damping, respectively, f_h represents the human operator input force, and u_s is generated using a proportional-derivative controller based on the delayed reference motion.

2) Delay Observer and Predictive Compensation: Communication delay is modelled as a time-varying disturbance affecting both forward and backward transmission channels. A first-order Taylor approximation of the delayed



(a) (b) (c)

Fig. 2: Representative simulation results under stochastic communication jitter. (a) Scenario 1: trapezoidal tracking in soft tissue. (b) Scenario 2: trapezoidal grasping with soft-to-hard contact transition. (c) Scenario 3: sinusoidal tracking at 1 Hz in free space.

master motion is defined as

$$x_m(t-\tau) \approx x_m(t) - \tau(t)\dot{x}_m(t), \tag{2}$$

where $\tau(t)$ denotes the unknown communication delay. An EKF is employed to estimate the delay online using delayed motion measurements and slave velocity information. The EKF state vector $\mathbf{x}_e = [\hat{x}_s,\ \hat{\dot{x}}_s,\ \hat{\tau}]^\top$ contains the estimated slave position, velocity, and communication delay. Based on the estimated delay, a predictive reference compensator reconstructs the current master reference according to

$$x_{\text{ref}} = k_m \left[x_m(t-\tau) + \hat{\tau}\dot{x}_m(t-\tau)\right]. \tag{3}$$

3) Environment Interaction Model: A switched contact model combining Kelvin–Boltzmann and Hunt–Crossley dynamics is implemented to emulate transitions between free-space motion, soft-tissue interaction, and hard contact conditions during robotic microsurgery.

The framework is evaluated in MATLAB/Simulink under a bounded stochastic delay (mean 0.50 s): correlated jitter ($T_c = 0.25$ s, $\sigma_\tau = 0.12$ s) combined with abrupt ±80 ms variations every 0.2 s, bounding total delay to 0.15–0.85 s. Scenarios 1 and 2 use trapezoidal motion profiles across free-space, soft-tissue, and switched contact conditions, while Scenario 3 uses sinusoidal excitation at 0.5, 1, and 2 Hz. Representative results are shown in Fig. 2.

Scenario 1 — trapezoidal-wave tracking. Under stochastic jitter, the EKF reduced delay RMSE from 126 ms to 98 ms in free space and to 104 ms in soft tissue (22% and 18% improvement, respectively). The uncompensated method exhibited the largest phase lag and tracking error. Delay observability is weakest during the dwell phases of the trapezoidal command, where $\dot{x}_m \approx 0$ (cf. Eq. 2); renewed motion excitation restores it and enables rapid EKF reconvergence.

Scenario 2 — grasping with environment switching. Across Kelvin–Boltzmann-to-Hunt–Crossley contact transitions, the proposed framework reduced force oscillations with no growing oscillatory behaviour, lowering delay RMSE from 127 ms to 118 ms and position RMSE from 4.82 mm to 4.52 mm.

Scenario 3 — sinusoidal frequency sweep. Continuous excitation aided low-frequency observability, cutting delay RMSE to 87 ms (0.5 Hz) and 105 ms (1 Hz) from a 126 ms baseline; at 2 Hz it rose to 140 ms (vs. 125 ms baseline), as rising latency variation and motion frequency exceed the observer–predictor bandwidth. This first-order predictor (Eq. 2) is computationally lightweight but neglects the higher-order motion terms that dominate under these conditions.

CONCLUSIONS AND DISCUSSION

The stochastic-delay simulations show that the proposed EKF improves online delay estimation under trapezoidal motion and low-frequency excitation, reducing delay RMSE by up to 23% relative to the fixed-gain baseline; however, improved estimation does not directly translate into better tracking or contact-force performance. Under hard-contact switching and high-frequency motion, residual estimation uncertainty and the first-order predictor limit compensation, defining an operating region for the approach rather than unconditional superiority under time-varying delay. These results demonstrate numerical closed-loop performance rather than formal stability or passivity. Experimental deployment on CARMA will additionally require force-sensor and vision-tracking noise, hardware/communication latency, RCM-constrained dynamics, and safety under delay-induced force transients; future work will address adaptive observer tuning and passivity-aware compensation alongside experimental validation on the physical platform.

REFERENCES

[1] N. S. Mastronikolis, et al "Management of laryngeal cancer: Current status and future perspectives," *European Archives of Oto-Rhino-Laryngology*, vol. 268, no. 1, pp. 5–14, 2011.

[2] S. Guo, et al "Bilateral teleoperation with time-varying delay using wave-variable compensation," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 5, pp. 2177–2188, 2019.

[3] Y. Wu, et al "Observer-based delay compensation for bilateral teleoperation systems under varying network latency," *IEEE Robotics and Automation Letters*, vol. 9, no. 2, pp. 1450–1457, 2024.

Inverse Design Framework for Tissue-Mimicking Phantoms: Applications in Computer and Robot-Assisted Surgery

Abu Bakar Dawood[1] Zhenyu Zhang[2] Dalia Osman[1] Cem Suulker[1]
Arsen Abdulali[1] Martin Angelmahr[2] Kaspar Althoefer[1]

[1]*School of Engineering and Materials Science, Queen Mary University of London, London, United Kingdom, E1 4NS*
[2] *Department of Fiber Optical Sensor Systems, Fraunhofer Heinrich Hertz Institute, Am Stollen 19H, 38640 Goslar, Germany*

INTRODUCTION

Computational models of soft tissues are essential for advancements in medical technologies including surgical simulation, robotic palpation, haptic feedback, and implantable device development. These models depend on constitutive parameters that describe the non-linear and nearly incompressible behaviour of soft tissues under compression, the dominant loading mode in surgical interactions [1]. However, biological tissue-derived parameters suffer from donor variability, tissue degradation, ethical constraints, and limited reproducibility, resulting in inconsistent mechanical predictions across studies[2], [3].

Hyperelastic constitutive models such as Neo-Hookean, Mooney-Rivlin, Yeoh, Ogden, Humphrey, and Veronda-Westmann are widely used to model soft tissues, but their accuracy depends on reliable parameter estimation [4]. Silicone phantoms provide a reproducible alternative through tunable stiffness, long-term stability, and repeatable testing. However, existing phantom datasets are predominantly tensile-based or do not adequately cover the stiffness range of soft tissues (5–150 kPa). Since hyperelastic parameters are loading-mode dependent, tensile-derived parameters inadequately represent compression-dominated surgical applications [5].

This work addresses this gap through an open-access database of compression-derived hyperelastic parameters from nineteen silicone phantoms spanning 6.19–158.80 kPa, covering tissues including adipose, liver, kidney, prostate, and skeletal muscle. The contributions of this extended abstract are:

1) compression-derived stress–strain datasets from ten repeated tests per phantom
2) fitted parameters for six hyperelastic constitutive models with statistical model selection metrics
3) inverse design framework linking target stiffness to phantom formulations
4) interactive organ browser mapping tissue targets to phantom compositions

All resources are openly accessible through the GitHub repository, including raw datasets, derived hyperelastic model parameters, and an inverse design framework linking target tissue elastic modulus to phantom fabrication recipes (github.com/abubakardawood/phantom-material-database) and archived on Zenodo [6].

MATERIALS AND METHODS

A. Phantom Fabrication

Nineteen cylindrical silicone phantoms (50 mm diameter × 30 mm height) were fabricated from EcoFlex 00-10, 00-30, and 00-50 (Smooth-On Inc.). Silicone thinner (0–100% by weight of Parts A+B) was mixed with Part A before adding Part B. Mixtures were degassed, cast in PLA moulds, and cured at room temperature [7]. Phantoms were

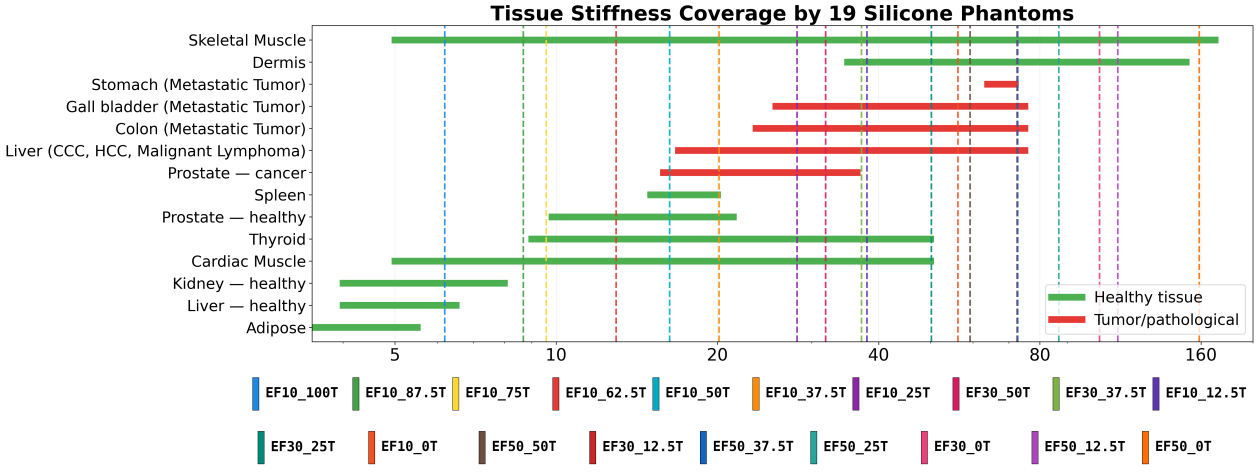


Fig. 1: Young's modulus ranges for healthy (green) and pathological (red) human tissues plotted on a logarithmic scale. Vertical dashed lines indicate the elastic moduli of the nineteen silicone phantoms, with the colour key shown below. The phantom database spans 6.19–158.80 kPa, covering tissues from adipose to skeletal muscle. [1], [2]

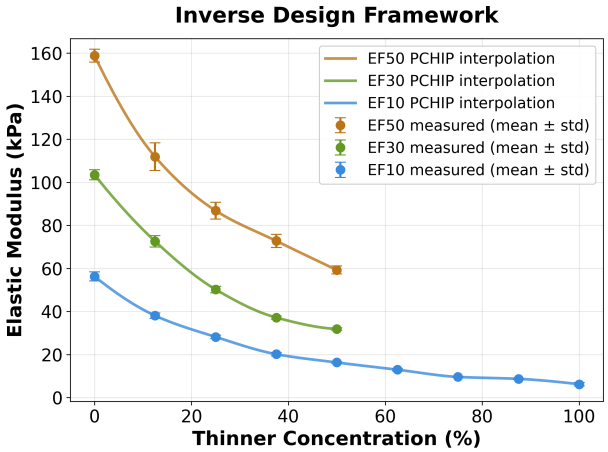


Fig. 2: Measured elastic modulus as a function of thinner concentration for the three silicone families.

labelled by silicone family and thinner concentration (e.g., $EF10_62.5T$ = EcoFlex 00-10 with 62.5% thinner).

B. Compression Testing

Compression tests were performed using an INSTRON 5900 with 1 kN and 10 kN load cells. Specimens were compressed to 20 mm displacement (66.6% strain) at 10 mm/min. Twelve tests were performed per phantom (six on each side), with the first cycle on each side discarded, yielding ten valid repetitions [8].

C. Data Processing and Model Fitting

Engineering stress and strain were calculated using the initial cross-sectional area ($A_o = \pi r^2$, $r = 25$ mm) and height ($H_o = 30$ mm). True stress and strain were derived from engineering values using standard conversion equations ($\varepsilon_{\text{true}} = \ln\left(1 + \varepsilon_{\text{eng}}\right)$ and $\sigma_{\text{true}} = \sigma_{\text{eng}}\left(1 + \varepsilon_{\text{eng}}\right)$). Elastic modulus was determined as the secant modulus over the initial 2 mm compression depth. Mean values, standard deviations, and coefficients of variation were computed across ten repetitions per phantom. Six constitutive models (Neo-Hookean, Mooney-Rivlin, Yeoh, Ogden, Veronda-Westmann, Humphrey) were fitted to both engineering and true stress-strain data using Levenberg–Marquardt least-squares optimisation.

D. Inverse Design Framework

Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) between measured data points enabled smooth mapping from target elastic modulus to silicone composition. The framework successfully generates recipes for target stiffness within the validated range. The relationship between the elastic moduli and thinner concentration is shown in Figure 2 with error bars representing the standard deviation of each measured phantom with coefficient of variation ranging from 1.25% to 12.44% and an average of 3.66%.

DISCUSSION AND CONCLUSIONS

This work presents an open-access library of nineteen tissue-mimicking silicone phantoms spanning elastic moduli from 6.19–158.80 kPa, covering physiologically relevant tissues including adipose, liver, kidney, prostate, and skeletal muscle. The library provides mechanically validated references for applications in surgical simulation, robotic palpation, haptic rendering, and device evaluation.

The database includes compression-derived engineering and true stress–strain datasets together with fitted hyperelastic material parameters for six constitutive models. Unlike biological tissues, which exhibit donor variability, degradation, and experimental constraints, the silicone phantoms demonstrated high repeatability, enabling reproducible mechanical characterisation and inter-study comparability. Furthermore, the compression-specific characterisation addresses limitations of existing tensile-derived datasets for applications dominated by compressive loading.

Beyond material characterisation, the inverse design framework enables direct mapping from target tissue stiffness to phantom fabrication recipes, reducing iterative phantom development and improving accessibility for non-specialist users. All datasets, hyperelastic model parameters, and inverse design tools are openly available through the GitHub repository archived on Zenodo [6].

Future work will extend the framework to additional loading modes and viscoelastic characterisation to further support computational and robot-assisted surgical applications.

ACKNOWLEDGMENTS

This work was carried out under the PALPABLE project, funded by UK Research and Innovation (UKRI) through the UK government's Horizon Europe funding guarantee [grant N°101092518] and funded by the European Union.

This work is supported by ERC grant EndoTheranostics, 101118626. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] G. Singh and A. Chanda, "Mechanical properties of whole-body soft human tissues: A review," *Biomedical Materials*, vol. 16, 2021.

[2] C. F. Guimarães, L. Gasperini, A. P. Marques, and R. L. Reis, "The stiffness of living tissues and its implications for tissue engineering," pp. 351–370, 5 2020.

[3] M. Griffin, Y. Premakumar, A. Seifalian, P. E. Butler, and M. Szarko, "Biomechanical characterization of human soft tissues using indentation and tensile testing," *Journal of Visualized Experiments*, vol. 2016, 2016.

[4] P. A. Martins, R. M. Jorge, and A. J. Ferreira, "A comparative study of several material models for prediction of hyperelastic properties: Application to silicone-rubber and soft tissues," *Strain*, vol. 42, 2006.

[5] L. Marechal, P. Balland, L. Lindenroth, F. Petrou, C. Kontovounisios, and F. Bello, "Toward a common framework and database of materials for soft robotics," *Soft Robotics*, vol. 8, pp. 284–297, 2021.

[6] A. B. Dawood and K. Althoefer, "A database of silicone phantom materials for modeling human tissue mechanics across varying stiffnesses," May 2026. [Online]. Available: https://doi.org/10.5281/zenodo.20262978

[7] D. G. Raitt, M. Huseynov, S. Homer-Vanniasinkam, H. A. Wurdemann, and S. A. Abad, "Soft-tipped sensor with compliance control for elasticity sensing and palpation," *IEEE Transactions on Robotics*, vol. 40, 2024.

[8] L. Paternò, G. Tortora, and A. Menciassi, "Hybrid soft-rigid actuators for minimally invasive surgery," *Soft Robotics*, vol. 5, 2018.

Bimanual Structure in Surgical Preparation

Eugenia Mawuenya Akpo[1,2], Nicholas Raison[2], Christos Bergeles[2], and Letizia Gionfrida[1,2]

[1]*Department of Informatics, King's College London, UK*

[2]*School of Biomedical Engineering & Imaging Sciences, King's College London, UK*

INTRODUCTION

Surgical preparation is a structured yet overlooked component of the perioperative workflow [1]. Recent work has begun exploring robotic systems for circulating nurse tasks, such as automated sterile package handling [2]. Despite its repetitive and rule-governed nature, surgical preparation remains largely absent from robotic automation pipelines. This gap is not solely due to task complexity, but to the lack of representations that translate human bimanual behaviour into robot-executable structure [3].

First-person (egocentric) vision-based Learning from Demonstration (LfD) provides a pathway to bridge this gap by learning directly from human behaviour [3]. However, progress is limited by two factors: the absence of structured representations linking first-person perception to manipulation primitives, and the lack of annotated datasets capturing bimanual coordination under realistic operating room (OR) conditions, including handling transparent and deformable packaging, self-occlusions, and specular reflections [4].

We aim to address some of these limitations by introducing a publicly available first-person RGB-D dataset collected in an OR and a structured bimanual workflow representation linking first-person vision to manipulation primitives via phases, affordances, and events. We show that surgical preparation exhibits consistent canonical structure (98% adherence), strong bimanual coordination (0.84 synchrony), and stable, object-dependent affordances.

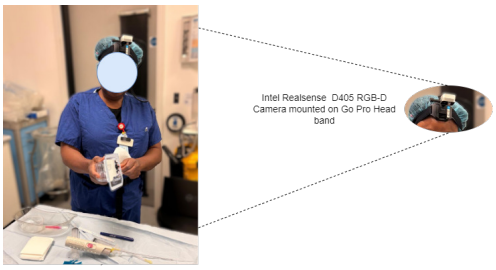


Fig. 1: First-person data acquisition: nurse wearing head-mounted Intel RealSense D405 camera.

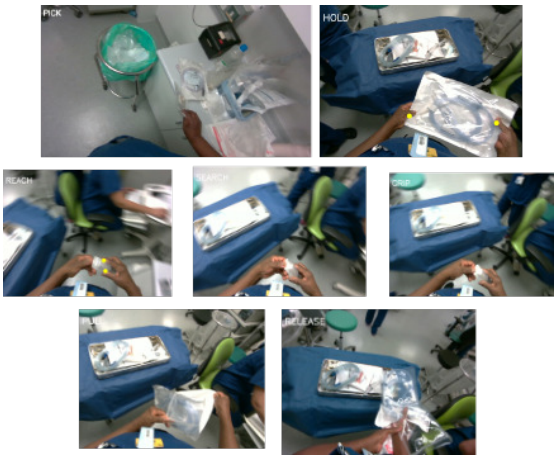


Fig. 2: The seven bimanual manipulation phases: pick, hold, reach, search, grip, pull, release (reading left to right, top to bottom). Hand landmarks are shown in green (left hand) and blue (right hand).

MATERIALS AND METHODS

Dataset

The collected dataset in an OR comprises **77 first-person RGB-D sequences** from **seven OR nurses** performing sterile package opening tasks in realistic scenarios, including self-occlusions, deformable and transparent materials handling. Data were captured using a head-mounted Intel RealSense D405 at 30 Hz. The dataset is publicly available at repository.

From this dataset, **25 sequences from five nurses** were fully annotated with phases, affordances, and task events for workflow analysis. Annotation was performed using a custom four-pass graphical tool: Pass 1 and 2 annotated left and right hand phases (pick, hold, reach, search, grip, pull, release) by marking start and end frames; Pass 3 marked spatial affordance points (grasp points, peel tabs, seal edges) via mouse clicks on keyframes; Pass 4 recorded temporal events (seal break, contents drop) at specific frames.

Workflow Representation

We model surgical preparation using a four-pass annotation framework:

Phases: Seven manipulation phases (pick, hold, reach, search, grip, pull, release) annotated per hand over time.

Affordances: Spatial interaction points including grasp regions, peel tabs, and seal edges.

Events: Task-level markers such as seal break and content release.

Pull synchrony is computed per sequence as the intersection-over-union of the left- and right-hand pull-phase frame intervals, $S = |I_L \cap I_R|/|I_L \cup I_R|$, averaged across sequences where both hands pull.

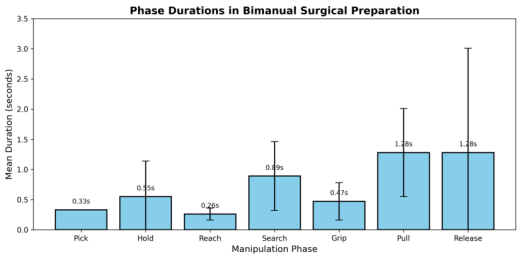


Fig. 3: Mean phase durations with standard deviation. Release shows high variance due to different nurse strategies for handling opened packages.

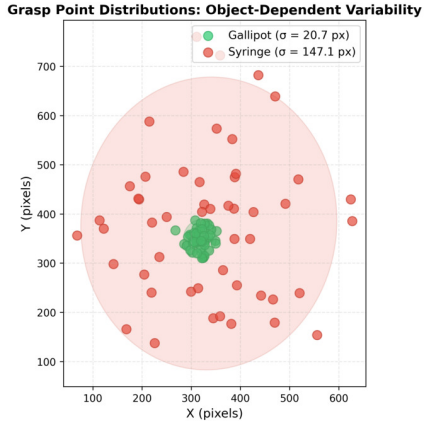


Fig. 4: Grasp point distributions: gallipot (tight, σ = 20.7px) versus syringe (wide, σ = 147.1px).

Hand Tracking and Refinement

Hand poses are estimated using MediaPipe [5] and refined using a spatio-temporal graph neural network [4]. Robot feasibility is evaluated using MuJoCo simulation with a Shadow Hand model on a subset of poses.

RESULTS

a) Workflow Regularity: The task follows a consistent canonical phase sequence in 98% of sequences. Transitions are near-deterministic, with $P(\text{search} \rightarrow \text{grip}) = 1.00$ and $P(\text{grip} \rightarrow \text{pull}) = 0.98$. The mean pull duration is 1.28 ± 0.73 seconds.

b) Bimanual Coordination: Bimanual interaction exhibits strong temporal coupling, with a mean pull synchrony of 0.84 (see Bimanual Synchrony Metric, above). Role asymmetry is consistent, with the right hand stabilising and the left hand manipulating in 60% of sequences.

c) Affordance Structure: Grasp distributions depend on object geometry. Compact objects exhibit tight clustering, while elongated objects show broader distributions. Intra-user variability is 55% lower than inter-user variability, indicating consistent individual strategies.

d) Robot Feasibility: Graph-based refinement reduces inverse kinematics error by 20.8%, with negligible joint limit violations, indicating the extracted workflow can be translated into physically feasible robotic motion.

TABLE I: Summary of key results.

Metric	Value
Annotated sequences	25 (5 nurses)
Canonical order adherence	98%
Pull duration	1.28 ± 0.73 s
Pull synchrony	0.84
Right stabilises, left manipulates	60%
Intra-user variance	σ = 74.1 px
Inter-user variance	σ = 115.1 px
IK error reduction	20.8%
Joint violations	< 0.04%

CONCLUSIONS AND DISCUSSION

This work demonstrates that surgical preparation exhibits consistent bimanual structure, making it suitable for automation. The near-deterministic phase transitions and high pull synchrony indicate that complex manipulation can be decomposed into stable, interpretable primitives. The high variance in release duration reflects different nurse strategies; some discard the package immediately, while others hold it to maintain sterility, suggesting internal substructure to explore in future work. Grasp consistency is object-dependent, and lower intra-user than inter-user variability points to consistent individual strategies with implications for personalised robotic assistance.

The representation is intended to map onto robotic control: phases select discrete motion primitives (reach, grip, pull), affordances supply the target poses and contact points that parameterise each primitive, and events trigger transitions between them. The feasibility results above are a first step toward closing this loop with a learned policy.

Limitations include the relatively small dataset size, simulation-only validation, and a primarily descriptive evaluation. Future work will scale the dataset, evaluate downstream tasks (phase/action recognition, a preliminary imitation-learning policy), integrate the workflow model with real-time robotic control on hardware, and formalise the affordance set as a lightweight ontology (object class, affordance type, grasp geometry) coupled to the phase sequence as a temporal layer.

Ethical approval: Approved by the King's College London Research Ethics Committee (MRSP-24/25-51732). All participants provided informed consent.

REFERENCES

[1] H.-J. Jang and S.-M. Lee, "Influence of surgery preparation time on patient outcomes," *Journal of Nursing Management*, vol. 2024, no. 1, p. 6753210, 2024.

[2] L. Bernhard, O. Baumann, C. Scheppach, P. Schwingenschlögl, C. Kössler, W. Haag, C. Müller, H. Feussner, A. Knoll, and D. Wilhelm, "The robotic circulating nurse," *IEEE Robotics & Automation Magazine*, 2025.

[3] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[4] J. Liu, A. Andres, Y. Jiang, Y. Du, X. Luo, W. Shu, C. Pu, and S. Tsaftaris, "Towards surgical task automation: Actor-critic models meet self-supervised imitation learning," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2025, pp. 7981–7988.

[5] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.-L. Chang, and M. Grundmann, "Mediapipe hands: On-device real-time hand tracking," *arXiv preprint arXiv:2006.10214*, 2020.

Evaluation of Foundation Models for Predicting Visual Outcomes after Epiretinal Membrane Surgery

Giulia Galvan[1,2,*], Ivo De Clerck[3,*], Mouloud Ourak[1], Elena De Momi[2], Emmanuel Vander Poorten[1]
[1]Department of Mechanical Engineering, KU Leuven, Leuven, Belgium
[2]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[3]Department of Ophthalmology, University Hospitals Leuven, Belgium

INTRODUCTION

Epiretinal membrane (ERM) is a retinal disorder characterized by the formation of a fibrous layer on the retinal surface, leading to visual distortion and reduced visual acuity. Although surgical removal is the standard treatment, predicting postoperative visual acuity (VA), a functional measure of the ability to resolve fine spatial details and typically expressed in logMAR units, remains challenging and often relies on clinical experience [1]. Accurate preoperative prediction could support clinical decision-making and also contribute to the development of robot-assisted ERM peeling procedures, by enabling more informed surgical planning and more adaptive, patient-specific robotic strategies [2]. Recent advances in artificial intelligence have shown promising results in ophthalmology, particularly for image-based analysis of Optical Coherence Tomography (OCT) scans. Early applications of deep learning focused primarily on automated disease detection [3], followed by approaches aimed at quantifying structural retinal features [4]. More recent studies have addressed postoperative outcome prediction both as classification and regression problems, enabling a more precise and clinically relevant estimation of patient outcomes [5], [6], [7]. Convolutional neural networks (CNNs), including architectures such as ResNet, have demonstrated strong predictive performance, with reported accuracies up to 80–87%. In parallel, increasing attention has been given to explainable artificial intelligence (XAI), which provides visual explanations of model predictions and helps identify retinal regions influencing the outcome, thereby improving transparency and clinician trust [8]. More recently, foundation models pretrained on large-scale retinal datasets have emerged as a promising alternative, enabling improved generalization and data efficiency [9]. For this reason, this work investigates the use of a retinal foundation model (RETFound) for postoperative VA prediction, alongside conventional deep learning architectures as baselines [10].

MATERIALS AND METHODS

Different predictive models were evaluated. First, baseline machine learning models using only clinical variables (e.g., age, sex, lens status) were implemented using Random Forest and multilayer perceptrons (MLP). Second, a CNN based on ResNet101 was trained using OCT images. The network was initialised with pre-trained weights on ImageNet, using transfer learning as a starting point for feature extraction. The original classification head with 1000 outputs was removed and replaced with a regression head combining image features with preoperative VA. A multimodal architecture combining OCT features with clinical data was also explored. Tabular features were processed through an MLP to obtain a compact 32-dimensional embedding, which was then fused with the visual features extracted by ResNet101. Finally, the RETFound retinal foundation model was fine-tuned for VA prediction. Pretrained on more than 1.6 million unlabeled retinal images [9], the encoder was adapted to preoperative OCT scans of ERM patients and connected to a regression head implemented as a simple MLP to predict postoperative VA in logMAR. The dataset consisted of OCT scans, acquired using a Zeiss Cirrus OCT device, and clinical data from patients who underwent ERM surgery at University Hospitals Leuven between 2015 and 2023, under ethical approval (study number S68853). Only the fovea-centered scan was selected for the analysis. In total 845 patients were retained for the 2-month postoperative prediction task and 419 for the 6-month task, due to the availability of follow-up data at the respective time points. The dataset was randomly divided into three subsets for model development: 80% for training, 10% for validation, and 10% for testing. All the images were resized (224×224 pixels) and data augmentation strategies were implemented during the training phase, including horizontal flipping and small-angle rotations (±10°). Model performance was evaluated using mean absolute error (MAE), defined as $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$, where y_i and $\hat{y}_i$ denote the ground truth and predicted values, respectively. Along similar lines, the root mean squared error (RMSE) and the coefficient of determination (R^2) were computed to provide complementary information on prediction accuracy and error distribution, enabling comparison with previously reported studies [7]. XAI analysis was performed using Grad-CAM that shows a localisation map that highlights the regions that are positively contributing to the prediction [8]. To provide an additional level of validation, the Grad-CAM heatmaps (such as depicted in Fig.1) generated by the network were qualitatively assessed by a retinal specialist, ensuring that the model's focus aligned with clinically relevant areas and reasoning patterns.

*These authors contributed equally to this work.

RESULTS

Model	MAE	MSE	RMSE	R^2
MLP (clinical)	0.1229	0.0299	0.1730	-0.0568
Random Forest (clinical)	0.1052	0.0236	0.1537	0.1666
RETFound (2 months)	0.0310	0.0023	0.0478	0.9244
ResNet101 (2 months)	0.0480	0.0043	0.0680	0.8541
RETFound (6 months)	0.0345	0.0029	0.0535	0.8834
ResNet101 (6 months)	0.0530	0.0068	0.0826	0.7463
ResNet101 + clinical	0.0544	0.0067	0.0817	0.7704

TABLE I: Model's performance comparison

Table I summarizes the results. The models trained only on the clinical variables showed limited predictive performance, with the best result obtained using Random Forest (MAE=0.1052, R^2=0.1666). In contrast, image-based models significantly improved accuracy, with RETFound consistently outperforming all others ($p<0.05$). In particular, the RETFound model achieved the best performance on the 2-month dataset (MAE=0.0310, R^2=0.9244), while ResNet101 also showed strong results (R^2=0.8541). For the 6-month dataset, RETFound maintained high performance (R^2=0.8834), while ResNet101 showed a decrease (R^2=0.7463). Integrating clinical data did not significantly improve performance (R^2=0.7704, $p>0.05$), suggesting that most relevant information is captured by OCT images. ResNet101 tended to underestimate BCVA at higher values (>0.6 logMAR), which may lead to overly optimistic predictions in clinical practice. Grad-CAM visualizations confirmed that models focused on clinically relevant retinal structures, including the foveal pit, ectopic inner foveal layers, and ellipsoid zone (Fig. 1).

CONCLUSIONS AND DISCUSSION

The ResNet models achieved performance comparable to previous studies (80–87%), while avoiding manual segmentation or handcrafted feature extraction, improving clinical applicability and reducing pre-processing complexity. The strong performance obtained using only the fovea-centered scan further suggests that relevant prognostic information can be effectively captured without processing the full OCT volume, reducing redundancy and noise. The RETFound model consistently outperformed conventional CNN architectures, highlighting the importance of domain-specific pretraining in medical imaging and transformer based architecture. This result supports the hypothesis that large-scale retinal pretraining enables the extraction of more robust and generalizable features. In addition, both the models based on clinical informations and and the multimodal model indicate that the majority of predictive information is contained within OCT images, while tabular clinical data provide limited additional benefit. The slight decrease in performance observed for the 6-month dataset may be attributed to the reduced sample size and to the increased complexity of long-term visual recovery, which is likely influenced by additional factors not captured in the current dataset. Moreover, the observed underestimation of high VA values suggests a potential limitation of the model in handling extreme cases, which may have implications for clinical interpretation, but these are the most relevant cases to predict because visual recovery is more variable in advanced stages of the disease [11]. Overall, this work demonstrates that deep learning models, and in particular foundation models, can accurately predict postoperative visual outcomes after ERM surgery and these results also support the integration of predictive models into robot-assisted retinal surgery, where outcome estimation could enhance surgical planning and personalization. The consistency between Grad-CAM visualizations and known anatomical biomarkers supports the biological plausibility of the predictions and may increase clinician trust in AI-based decision support systems. Future work will focus on multi-center external validation to further assess the generalizability and clinical applicability of the proposed approach in diverse patient populations and with scans obtained with different sensors.

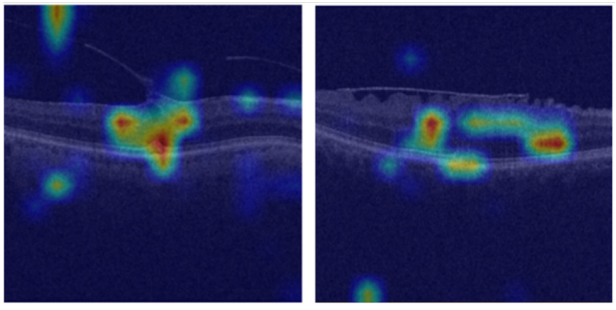


Fig. 1: Example of Grad-CAM visualizations highlighting relevant retinal regions contributing to model prediction. Warmer colors indicate regions with higher contribution.

REFERENCES

[1] Adrian T. Fung. Epiretinal membrane: A review. *Journal*, 2021.

[2] Ferhat Turgut. Robot-assisted epiretinal membrane peeling: A prospective assessment of pre- and intra-operative times and of surgeons' subjective perceptions. *Journal of Clinical Medicine*, 2023.

[3] Ying Chih Lo. Epiretinal membrane detection at the ophthalmologist level using deep learning of optical coherence tomography. *Scientific Reports*, 10, 2020.

[4] Soo Han Kim. Deep learning-based prediction of outcomes following non complicated epiretinal membrane surgery. *Technical Report*, 2022.

[5] Emanuele Crincoli. New artificial intelligence analysis for prediction of long-term visual improvement after epiretinal membrane surgery. *Technical Report*, 2023.

[6] Hsin Lin. Using a deep learning model to predict postoperative visual outcomes of idiopathic epiretinal membrane surgery. *American Journal of Ophthalmology*, 272, 2025.

[7] Dejia Wen. Deep learning-based postoperative visual acuity prediction in idiopathic epiretinal membrane. *BMC Ophthalmology*, 23, 2023.

[8] Ramprasaath R Selvaraju. Grad-cam: Visual explanations from deep networks via gradient-based localization. *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[9] Yukun Zhou. A foundation model for generalizable disease detection from retinal images. *Nature*, 622, 2023.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 2015.

[11] Ivo De Clerck. Surgical outcomes, validation of govetto staging, and postsurgical macular edema in idiopathic epiretinal membranes: A large retrospective study. *Retina*, 2025.

Vision-Based Hand-Motion Analysis for Ergonomic Assessment in Emulated Robot-Assisted Partial Nephrectomy

Deborah Cattafesta[1*], Elisa Iovene[1], Mariachiara Di Cosmo[2], Federica Ferraguti[1], Sara Moccia[2]

[1]*Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Italy*
[2]*Department of Innovative Technologies in Medicine and Dentistry, Università degli Studi "G. d'Annunzio" di Chieti-Pescara, Italy*
**deborah.cattafesta@unimore.it*

INTRODUCTION

Robot-assisted surgery (RAS) is widely used for complex minimally invasive procedures such as Robot-Assisted Partial Nephrectomy (RAPN), due to its advantages in dexterity, precision, and visualization. However, robotic console use may still impose physical demands on the operator, with prolonged tasks contributing to fatigue, discomfort, and musculoskeletal strain [1].

Quantitative ergonomics assessment is gaining increasing interest in RAS. Objective measurements of hand motion and body posture can provide useful insights into how users interact with the robotic console, how motor strategies vary with task complexity, and whether different operators adopt more or less efficient behaviors [1].

Since hand and finger movements directly mediate console manipulation, their quantitative analysis can characterize operator-side behavior that is not fully represented by instrument kinematics. Robot kinematics describe the motion transmitted to the robotic system, whereas vision-based hand tracking directly captures hand behavior at the console and may therefore provide complementary information for ergonomic assessment, as clutching, motion scaling, and static holding phases can decouple hand movements from instrument trajectories [2]. Markerless solutions such as MediaPipe Hands [3] can estimate hand landmarks without dedicated wearable sensors; however, their reliability in the constrained and occluded environment of a surgical console has not been systematically evaluated.

This work investigates hand-motion behavior during emulated RAPN performed on the da Vinci Research Kit (dVRK) [4]. We evaluate the agreement between MediaPipe Hands predictions and manual annotations across surgical phases and task difficulty levels, and extract kinematic descriptors for subsequent ergonomic analysis of operator–console interaction.

MATERIALS AND METHODS

Experimental setup. Experiments were conducted on the dVRK platform [4], comprising the surgeon console and the patient-side unit equipped with two Patient Side Manipulators and one Endoscopic Camera Manipulator. An emulated RAPN task was designed to reproduce four key surgical phases: renal artery clamping, tumor resection, single-stitch suturing, and tumor extraction. Three levels of task difficulty (L1–L3) were defined based on increasing the anatomical complexity (i.e., number of vessels and tumor location and visibility) of a patient-specific silicone phantom reconstructed from CT data. Nine participants were enrolled (mean age: 24.33 ± 1.80 years; 8 right-handed and 1 left-handed; 5 males and 4 females), none with prior laparoscopic or robotic surgery experience; 6 had previous exposure to the da Vinci platform through phantom-based simulation. All participants provided written informed consent, and the study protocol (n.17/2026) was approved by the Ethics Committee of Politecnico di Milano.

Data acquisition and processing. Hand motion and upper-body posture were recorded using three Intel RealSense RGB-D cameras: two D435i devices mounted inside the surgeon console for hand tracking and one lateral D455i camera for posture monitoring. Data acquisition was managed using ROS Noetic on Ubuntu 20.04. The robot operated at 500 Hz, while cameras recorded at 15 fps. For each surgical phase, 15 representative frames were selected and manually annotated. Model predictions were obtained by processing the same frames with MediaPipe Hands [3], a two-stage framework that first detects the palm and then regresses 21 2.5D hand landmarks from the cropped hand region. Frames with absent hand motion were excluded from the analysis, as non-informative for the assessment of dynamic ergonomic behavior.

Hand-motion evaluation metrics. Frame-to-frame Euclidean distance (in pixels, px), speed (px/s), and acceleration (px/s^2) were computed from tracked landmarks for both model predictions and manual annotations. Agreement between predicted and annotated metrics was assessed using Spearman's ρ and Pearson's r correlation coefficients [5], and analyzed across surgical phases and difficulty levels.

RESULTS

Figure 1 shows a representative example of hand landmark estimation during task execution. Overall, MediaPipe Hands predictions show good spatial correspondence with manual annotations, particularly in the central region of the hand. Larger discrepancies are visible for peripheral landmarks, especially distal finger keypoints.



Fig. 1: Representative frame acquired from the internal RGB-D camera during task execution. Colored circles indicate manual annotations; white crosses represent the corresponding MediaPipe Hands predictions.

TABLE I: Correlation between motion-derived metrics computed from MediaPipe Hands predictions and manual annotations across RAPN phases, in terms of Spearman's ρ / Pearson's r.

Metric	Clamping	Resection	Suturing	Tumor ext.
Eucl. dist.	0.67 / 0.76	0.82 / 0.87	0.77 / 0.82	0.64 / 0.66
Speed	0.66 / 0.75	0.83 / 0.89	0.82 / 0.88	0.68 / 0.66
Acceleration	0.81 / 0.85	0.86 / 0.89	0.81 / 0.83	0.75 / 0.80

TABLE II: Motion-metric agreement across RAPN difficulty levels. Columns represent increasing task complexity (L1-L3), and cells report Spearman's ρ / Pearson's r.

Metric	L1	L2	L3
Eucl. dist.	0.82 / 0.81	0.82 / 0.82	0.56 / 0.72
Speed	0.76 / 0.82	0.85 / 0.85	0.60 / 0.75
Acceleration	0.84 / 0.83	0.86 / 0.88	0.85 / 0.89

As shown in Table I, agreement varied across RAPN phases. Higher correlations were observed during resection and suturing, where Euclidean distance reached $\rho/r = 0.82/0.87$ and $0.77/0.82$, respectively. Lower values were found during clamping and tumor extraction, with the lowest Euclidean distance agreement in tumor extraction ($0.64/0.66$). Acceleration showed the most consistent agreement across phases.

The level-wise analysis reported in Table II showed moderate-to-high agreement at L1 and L2. At L3, Euclidean distance and speed decreased to $\rho/r = 0.56/0.72$ and $0.60/0.75$, respectively, suggesting reduced reliability under higher task complexity. In contrast, acceleration remained stable across levels ($0.84/0.83$–$0.85/0.89$).

DISCUSSION AND CONCLUSIONS

The results support the feasibility of vision-based hand-motion analysis for ergonomic assessment during emulated RAPN. Across phases, the mean participant-level percentage of non-visible annotated keypoints ranged from 39.65% to 40.81%, with the highest mean proportion observed during clamping (40.81%). Occlusions mainly affected thumb joints and distal fingertips, likely due to finger overlap and partial hiding caused by the grasping posture on the robotic manipulators, contributing to the lower agreement observed in this phase.

The lower agreement during clamping may also reflect its spatial characteristics: compared to other phases, clamping showed a more limited horizontal excursion with the smallest median range across keypoints (266 px vs. 351 px in resection), combined with the highest dispersion of central keypoint positions (median interquartile range: 133 px vs. 115 px in resection). This suggests that compact but variable hand configurations may reduce tracking reliability. Across difficulty levels, displacement and speed showed comparable or higher correlations at L2 than at L1, but decreased at L3, particularly for Spearman's correlation. This suggests that the most complex condition introduced greater hand-motion variability and less stable landmark tracking. The higher agreement at L2 may reflect a familiarization effect, as the first execution likely served as an adaptation phase for participants with limited experience. Conversely, the stability of acceleration across levels suggests that this metric may be less sensitive to difficulty-related variability and could provide a more robust descriptor of hand-motion dynamics.

Overall, this study presents a preliminary vision-based framework for quantitative hand-motion analysis during emulated RAPN. The observed agreement across phases and difficulty levels supports the feasibility of extracting operator-side motion descriptors in the constrained and partially occluded console environment, thus complementing kinematics by directly characterizing hand behavior at the console. Future work will compare hand-motion estimates with master-manipulator kinematics, integrate hand and posture measures into established ergonomic indices, and extend validation to larger clinical cohorts.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Italian Ministry of University and Research under the D.D. 1236 (August 1, 2023), Fondo Italiano per la Scienza – Bando FIS 2 (TRAMIS – Trustworthy Robotic Assistant for Improved Minimally Invasive Surgery, CUP E53C24003820001).

REFERENCES

[1] Wong, S. W., Parkes, A., and Crowe, P. Ergonomic interventions to reduce upper limb musculoskeletal pain during robotic surgery: a narrative review. *Journal of Robotic Surgery*, 18(1):224, 2024.

[2] Hashemi, N., Svendsen, M. B. S., Bjerrum, F., Rasmussen, S., Tolsgaard, M. G., and Friis, M. L. Acquisition and usage of robotic surgical data for machine learning analysis. *Surgical Endoscopy*, 37(8):6588–6601, 2023.

[3] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.-L., and Grundmann, M. MediaPipe Hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214*, 2020.

[4] Kazanzides, P., Chen, Z., Deguet, A., Fischer, G. S., Taylor, R. H., and DiMaio, S. P. An open-source research kit for the da Vinci® Surgical System. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6434–6439, 2014.

[5] Schober, P., Boer, C., and Schwarte, L. A. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5):1763–1768, 2018.

Towards Sensor-Free Surgical Skill Assessment Using Video-Based Multimodal Learning

Ali Hesham Osman, Duygu Sarikaya

School of Computer Science, University of Leeds
sc23ao@leeds.ac.uk

INTRODUCTION

Surgical care is a vital component of modern healthcare systems, with approximately 800,000 procedures performed every day globally [1]. Technical errors account for 50–75% of surgical complications, underscoring the critical role of rigorous, objective training and skill assessment [2]. Traditional evaluation relies on the Halstedian apprenticeship model, "see one, do one, teach one", a process that is time-consuming, expensive, and prone to inter-rater bias. There is therefore an urgent need for automated, data-driven assessment systems capable of delivering consistent, actionable feedback at scale, particularly given a global shortage that leaves an estimated 5 billion people without access to safe surgical care [3].

Most existing automated approaches frame skill assessment as a three-class classification problem (novice, intermediate, expert); this discards fine-grained performance differences and provides little actionable feedback to trainees. In contrast, real-world assessment requires predicting a continuous score like the Global Rating Scale (GRS) that reflects subtle variations in performance and is clinically more informative. Existing multimodal approaches combine kinematic data with video; however, their reliance on specialised robotic sensors limits scalability beyond robot-assisted surgery. Consequently, there is a need for approaches that retain the benefits of multimodal models while removing dependency on privileged sensor data. To address these challenges, we propose a sensor-free deep learning framework for continuous GRS regression that combines visual data with 2D pose estimation as an alternative to privileged kinematic data.

MATERIALS AND METHODS

A. Dataset and Preprocessing

Experiments are conducted on the suturing task of the JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS) [4], comprising 39 trials from 8 right-handed surgeons (4 novices, 2 intermediates, 2 experts) recorded using the da Vinci Surgical System at 640×480 pixels and 30 FPS. GRS annotations span 6–30 points across six sub-criteria: Respect for Tissue (RT), Suture/Needle Handling (SH), Time and Motion (TM), Flow of Operation (FO), Overall Performance (OP), and Quality of Final Product (QP). 2D skeleton annotations were generated in-house using DeepLabCut, producing per-frame coordinates for 10 tool joints (5 per arm), as shown in Figure 1.

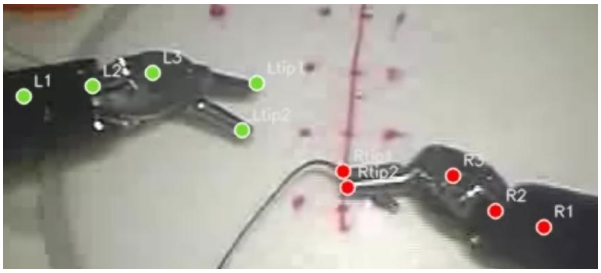


Figure. 1: 2D skeleton annotations overlaid on video

Joint tracks with confidence below 0.7 are interpolated; fully unreliable tracks are zero-filled. Six kinematic-proxy motion features are derived per joint: velocity, acceleration, jerk, inter-tip distance, angular velocity, and angular acceleration. These simulate economy-of-motion and smoothness signals otherwise provided by sensor data. The six motion features are concatenated with the joint's raw (x, y) coordinates and confidence score, yielding a 9-dimensional per-joint vector. Videos are segmented into overlapping windows (length = 250 frames, stride = 200 frames). Data augmentation, like random noise injection, spatial scaling, temporal masking, and intensity scaling, are applied only during training.

B. Architecture

The proposed dual-branch architecture (Figure 2) consists of five components. The vision branch processes frame sequences through a ResNet18 backbone with ImageNet pre-training; early layers are frozen, and only the final convolutional block is fine-tuned to balance domain adaptation with regularisation. Global Average Pooling (GAP) then collapses the spatial feature maps into per-frame vectors. The skeleton branch models tool trajectories using three stacked ST-GCN blocks [5]. Each block combines spatial graph convolutional encoding of cross-instrument tip connections and intra-arm joints with a temporal convolution (kernel size 9, padding 4). Graph features are aggregated across joints via GAP, aligning the skeleton representation with the vision branch for fusion. Cross-modal attention fuses the two branches using scaled dot-product attention, in which the skeleton representation queries the visual features to ground noisy 2D coordinates in rich contextual information. Attention is restricted to a local temporal window (±30 frames) to prevent mean-regression collapse caused by near uniform global attention weights in early phases of training. The fused representation is passed through a Multi-Stage Temporal Convolutional Network (MSTCN) with dilated convolutions that refine predictions across time, enforcing temporal coherence [6].

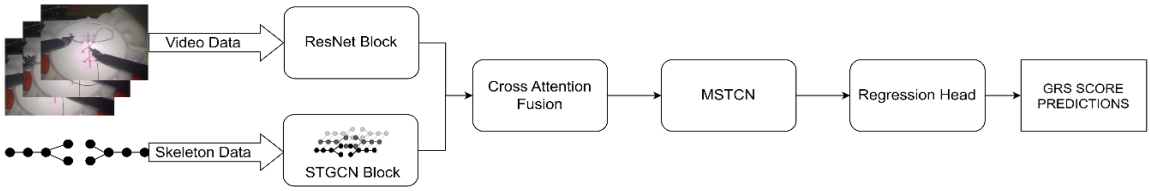


Figure. 2: Model Architecture

Finally, a convolutional multi-regression head jointly predicts the overall GRS and all six sub-scores, with the model trained using Adam at a learning rate of $1\ x\ 10^{-5}$, a batch size of 4, for up to 60 epochs on RTX 4060.

C. Training and Evaluation

A combined Rank-MSE loss penalises both absolute prediction error and incorrectly ordered surgeon pairs, encouraging clinically meaningful rankings. Three evaluation protocols are employed: Leave-One-Supertrial-Out (LOSO), Leave-One-User-Out (LOUO), and 4-fold cross-validation. Performance is measured using Spearman's rank correlation coefficient (ρ) as the primary metric to reflect the clinical priority of correctly ordering surgeons by skill and Mean Absolute Error (MAE) in GRS points.

RESULTS

Table 1 presents modality ablation results across all three protocols. The skeleton-only model achieves LOSO ρ = 0.573, confirming that 2D pose alone is insufficient, with tissue handling sub-scores (RT, FO) showing the highest errors (MAE ≈ +0.4) due to limited visual context. In contrast, Time and Motion is better captured by the skeleton modality. The vision-only model achieves ρ = 0.628 under LOSO, while the fused model outperforms both individual modalities in all protocols. Direct comparison with published JIGSAWS methods remains limited by differing protocols and modalities, with several prior studies not reporting LOUO results.

Under LOUO, performance degrades (ρ = −0.254 ± 0.544), with only 8 surgeons available for training, the ST-GCN backbone learns user-specific motion signatures (rhythm, trajectory smoothness) and usage patterns rather than generalisable skill representations. The cross-attention mechanism reinforces this by attending to consistent motion cues that act as proxies for identity rather than skill, causing rank inversion when evaluating unseen users whose motion styles differ substantially from those seen during training.

Modality	4-Fold	LOSO	LOUO
Vision	0.614 (3.69)	0.628 (3.32)	0.031 (5.64)
Skeleton	0.637 (4.73)	0.573 (4.80)	-0.078 (4.84)
Fused (Ours)	0.716 (3.53)	0.732 (3.34)	-0.254 (6.05)

Table. 1: Modality ablation results, Spearman ρ (MAE in GRS)

CONCLUSION AND DISCUSSION

We present a sensor-free multimodal framework for continuous surgical skill assessment that achieves competitive GRS regression without reliance on privileged robotic kinematic data. By integrating ResNet18 visual features with ST-GCN pose-derived motion representations through windowed cross-modal attention and MS-TCN temporal refinement, the model captures complementary skill signals that neither modality provides alone. This approach is applicable beyond robot-assisted environments; in open or laparoscopic surgery, pose estimation could be adapted to track the surgeon's hand key points, enabling similar assessment without specialist hardware. Given the small sample size, the absence of confidence intervals and significance testing warrants cautious interpretation of differences between modalities.

The principal limitation is the dataset scale: with 39 trials and 8 surgeons, JIGSAWS is insufficient to support cross-user generalisation of deep architectures, leading to identity entanglement and LOUO failure. Future work will pursue user-identity disentanglement through domain adaptation and adversarial training, alongside attention-map visualisation and per-trial failure analysis to improve interpretability.

REFERENCES

[1] Dobson, G.P. 2020. Trauma of major surgery: a global problem that is not going away. International Journal of Surgery. 81, pp.47–54.

[2] Fabri, P.J. and Zayas-Castro, J.L. 2008. Human error, not communication and systems, underlies surgical complications. Surgery. 144(4), pp.557–565.

[3] Meara, J.G., Leather, A.J.M., Hagander, L., et al. 2015. Global surgery 2030: evidence and solutions for achieving health, welfare, and economic development. The Lancet. 386(9993), pp.569–624.

[4] Gao, Y., Vedula, S.S., Reiley, C.E., et al. 2014. The JHUISI gesture and skill assessment working set (JIGSAWS): a surgical activity dataset for human motion modeling. Modeling and Monitoring of Computer Assisted Interventions (M2CAI) Workshop.

[5] Yan, S., Xiong, Y. and Lin, D. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In: Proceedings of the AAAI Conference on Artificial
Intelligence. 32(1).

[6] Abu Farha, Y. and Gall, J. 2019. MS-TCN: multi-stage temporal convolutional network for action segmentation.

Real-Time AI Generative Feedback for full-lung bronchoscopy instruction

Chuan Tian[1], Hao Yu[2], Haolin Jiang[1], Bruno Oliveira[1], Lars Konge[3], Kristoffer Cold[3], and Zhuoqi Cheng[1]

[1]*Maersk Mc-Kinney Moller Institute, SDU Robotics, University of Southern Denmark, Odense, Denmark*

[2]*Institute for Imaging, Data and Communications, University of Edinburgh, UK*

[3]*Copenhagen Academy for Medical Education and Simulation, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark*

INTRODUCTION

Achieving a complete and systematic examination of the bronchial tree is clinically important but technically demanding, particularly for novice bronchoscopists who must simultaneously learn bronchial anatomy, maintain spatial orientation within a highly branching lumen structure, and execute coordinated bronchoscope manipulations under continuously changing visual conditions. As many airway passages appear visually similar and anatomical landmarks can be difficult to distinguish, extensive training is required for trainees to master systematic full-lung bronchoscopy while maintaining smooth instrument handling and avoiding missed airway segments. Previous studies have shown that even trained physicians may fail to inspect all bronchial branches during bronchoscopy procedures, highlighting the intrinsic complexity of the task and the need for improved training support [1].

Currently, bronchoscopy training primarily follows a mentorship model, where novices acquire procedural skills through supervised practice using realistic phantoms or simulators under the guidance of experienced clinicians. Nevertheless, systematic reviews have reported persistent limitations in skill transferability, objective performance assessment, and real-time procedural feedback during training, motivating the development of intelligent systems capable of providing continuous and objective coaching. In fact, recent advances in artificial intelligence (AI) have further accelerated the development of computer-assisted bronchoscopy training systems. In particular, AI-based airway recognition and navigation systems have demonstrated the ability to provide real-time bronchial segment identification, navigation assistance, and automated performance analysis, thereby improving procedural training efficiency [2]. However, current AI systems remain largely passive guidance tools, with limited ability to explain procedural reasoning or respond to trainee uncertainty.

This motivates us to develop BronchoSAS, an AI-driven instructional agent for bronchoscopy navigation training. The system continuously analyzes the bronchoscopy scene and provides real-time operational recommendations, anatomical guidance, and emotional support during full-lung bronchoscopy. In addition, by incorporating natural language processing (NLP), trainees can interact directly with the system to ask questions, clarify procedural uncertainties, and receive contextual explanations during training.

MATERIALS AND METHODS

As shown in Fig. 1, BronchoSAS comprises runtime inputs, the BronchoSAS Core, post-processing, and live session outputs. AIBAN-derived navigation state, scope control state, and trainee voice feed the Core, which coordinates four real-time skills and session-level statistics and reporting.

The upstream module, AI Bronchial Detection and Airway Navigation (AIBAN) [3], is an AI-based framework to provide anatomical awareness and navigation guidance with real-time bronchoscopy video stream as input. This module outputs: (1) detection of visible airway branches, (2) estimation of the bronchoscope's anatomical position within the bronchial tree, and (3) recommendation of the next airway target direction.

The SAS Core is the central instructional module of BronchoSAS. It organizes upstream signals into a consistent internal state and interprets them through a procedure-specific curriculum and bronchoscopy knowledge package. The curriculum defines the inspection route and target sequence, while the knowledge package provides airway relationships, landmark cues, teaching cards, QA facts, and repair rules grounded in established systematic bronchoscopy and airway-anatomy sources [4], [5]. These resources structure runtime guidance and constrain it to predefined task knowledge rather than unrestricted model knowledge.

There are four real-time interaction skills to provide immediate instructional support during bronchoscope navigation. The guidance skill delivers moment-to-moment coaching toward the next airway target based on the current procedural state. The landmark teaching skill provides brief anatomy-focused teaching at key landmarks to support visual recognition and procedural orientation. The question-answering skill gives context-sensitive responses to learner questions without fully interrupting task progression. The support skill provides corrective or reassuring feedback when the learner appears disoriented or stuck.

Beyond real-time interactions, BronchoSAS also provides session-level assessment and reporting. The session report summarizes curriculum coverage, procedure time,

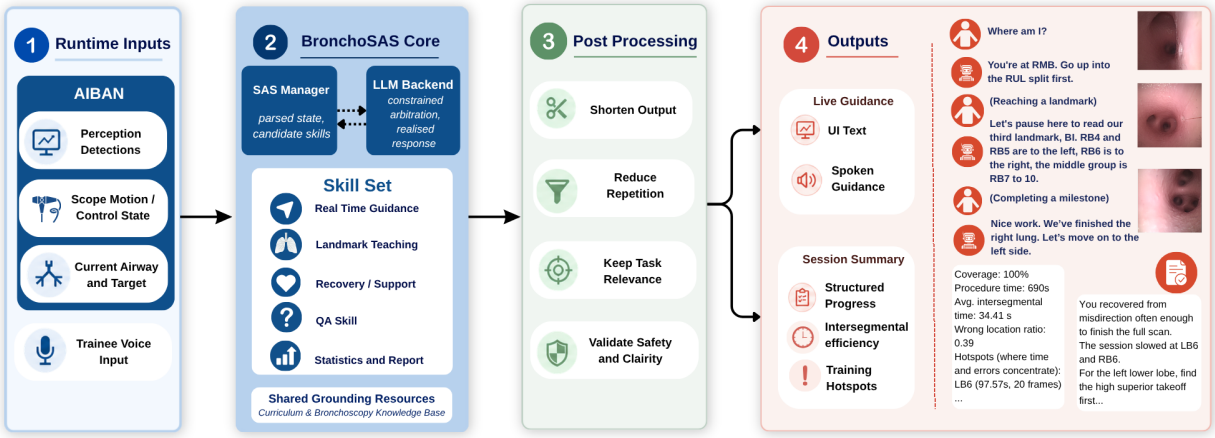


Fig. 1: The architecture of the BronchoSAS training system.

target-specific time hot spots, backtracking, and off-route events, allowing both overall performance and localized navigation difficulties to be reviewed.

To improve real-time usability while maintaining correctness, each selected skill first builds a grounded response frame from the procedural state, curriculum, and curated knowledge. The LLM only reformulates this frame into concise coaching language, but it cannot alter airway targets or the planned route. Post-processing checks key labels, actions, directional and safety intent, and falls back to deterministic wording if constraints are violated.

RESULTS

A single-condition formative evaluation was conducted to assess whether BronchoSAS provided understandable, usable, timely, and helpful support during phantom-based bronchoscopy training. Five novice users and one domain expert participated. After a brief familiarization, the novices completed a phantom-based navigation task with BronchoSAS enabled and subsequently completed an adapted BUS-15 questionnaire. The expert separately reviewed anatomical accuracy, procedural appropriateness, teaching usefulness, timing, safety, and conciseness.

The rightmost column of Figure 1 showcases the model's output including live guidance (upper) and session report (lower). In the live guidance outputs example, the texts inside bracket represent user's state, while those outside are conversation.Across the 15 adapted BUS-style items, the average item mean was 3.83/5. The strongest ratings concerned appropriate information amount (4.6/5), perceived response accuracy (4.4/5), and task relevance (4.2/5). Lower ratings concerned conversational continuity (2.8/5), waiting time (3.0/5), and handling of unclear task or dialogue states (3.2/5). These results suggest that BronchoSAS performed more strongly in context-aware task guidance than in conversational continuity and real-time interaction smoothness.

The expert review provided complementary domain-informed feedback. Safety and anatomical accuracy received the strongest ratings (5/5 and 4/5, respectively), while procedural appropriateness and conciseness were rated 3/5, and teaching usefulness and timing were rated 2/5. These ratings indicate that the main remaining issues concern when and how instructional support is delivered, rather than content accuracy alone.

Generated session reports were also inspected as formative outputs for consistency with logged measures such as completeness, structured progression, procedure time, efficiency, and locational accuracy. This was a development check rather than a validated clinical assessment. The reports additionally summarized strengths, risks, and next-session priorities.

CONCLUSION AND DISCUSSION

This study presents BronchoSAS, an AI generative natural language instructions from raw bronchoscope image sequences.Preliminary results suggest that the system is understandable and task-relevant, while timing and conversational continuity require further improvement. Future work will evaluate the approach in larger studies.

ACKNOWLEDGEMENT

This work was supported by Fabrikant Vilhelm Pedersen og Hustrus Legat.

REFERENCES

[1] K. M. Cold, K. O. Agbontaen, A. O. Nielsen, C. S. Andersen, S. Singh, and L. Konge, "Artificial intelligence for automatic and objective assessment of competencies in flexible bronchoscopy," *Journal of Thoracic Disease*, vol. 16, no. 9, pp. 5718–5726, 2024.

[2] K. O. Agbontaen, K. M. Cold, D. Woods, V. Grover, H. S. Aboumarie, S. Kaul, L. Konge, and S. Singh, "Artificial intelligence guided bronchoscopy is superior to human expert instruction for the performance of critical-care physicians: A randomized controlled trial," *Critical Care Medicine*, 2025.

[3] B. Oliveira, Z. Cheng, L. H. Guldbrøn, L. Konge, T. R. Savarimuthu, and K. M. Cold, "Aiban–ai bronchial detection and airway navigation system," *Journal of Bronchology & Interventional Pulmonology*, vol. 33, no. 3, p. e1073, 2026.

[4] K. M. Cold, A. Vamadevan, A. O. Nielsen, L. Konge, and P. F. Clementsen, "Systematic bronchoscopy: The four landmarks approach," *Journal of Visualized Experiments*, no. 196, 2023.

[5] B. Ryan, K. Yendamuri, and S. Yendamuri, "Anatomical considerations in bronchoscopy," *Journal of Thoracic Disease*, vol. 9, no. Suppl. 10, pp. S1123–S1127, 2017.

Real-Time Instance Segmentation in the Operating Room with Selective Refinement of Rare Surgical Instruments

R. Damiano[1], A. Ivanaj[1], O. Alfieri[2,3], P. Denti[2,3],
F. Maisano[2,3], A. C. L. Redaelli[1], E. De Momi[1], A. Marzullo[1]

[1]Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[2]Department of Cardiac Surgery, Vita-Salute San Raffaele University, IRCCS Ospedale San Raffaele, Milan, Italy
[3]Heart Valve Center, Vita-Salute San Raffaele University, IRCCS Ospedale San Raffaele, Milan, Italy
rossella.damiano@polimi.it

INTRODUCTION

The operating room (OR) is a highly complex and dynamic environment where clinicians and devices interact simultaneously. In intelligent ORs, ceiling-mounted cameras support global scene understanding for context-aware assistance and autonomous surgical systems. In this context, accurate instrument segmentation is a key component, providing pixel-level information about object boundaries, shapes, and spatial relationships that can support hand–instrument interaction analysis and downstream tasks such as workflow monitoring and robotic assistance. However, surgical scenes remain challenging due to visual variability and frequent occlusions. Small and rare instruments, used only during specific procedural steps and thus appearing in a small fraction of frames, are particularly difficult to segment because of limited visibility and underrepresentation in the data. Several deep learning methods have been developed for real-time instance segmentation. You Only Look Once (YOLO)-based approaches provide an effective trade-off between accuracy and inference speed, while Segment Anything Model (SAM)-based methods can produce high-quality segmentation masks, either automatically or guided by prompts such as points or bounding boxes. However, SAM is computationally demanding and does not inherently assign semantic class labels to the generated masks, requiring an additional classification step.
In this work, we propose a low-latency hybrid framework for OR tool instance segmentation, combining YOLO26 [1] with EfficientSAM [2], a lightweight SAM variant. YOLO26, fine-tuned on the 4D-OR dataset [3], provides real-time object localization, class predictions, and segmentation masks. EfficientSAM selectively refines masks of challenging underrepresented instruments using YOLO-predicted bounding boxes as spatial prompts. The resulting framework improves segmentation quality for rare surgical instruments while preserving real-time inference capability.

MATERIALS AND METHODS

Dataset

This study relies on the publicly available 4D-OR dataset [3], a multi-view collection of ten simulated total knee replacement interventions acquired in a realistic OR setting. The dataset includes 6734 scenes captured by six calibrated RGB-D sensors mounted on the ceiling, sampled at 1 frame per second (FPS). All experiments use *Camera #2* exclusively, which provides a consistent lateral viewpoint of the operative field and has complete panoptic annotation coverage. This single-camera setup limits viewpoint variability and facilitates a controlled assessment of segmentation performance. The provided annotations are used as ground truth for training and evaluation. Class frequency analysis over all 6734 frames revealed a strong class imbalance: three categories (*instrument table*, *secondary table*, *drape*) appear in over 67% of frames, whereas four surgical tool categories (*hammer*, *saw*, *drill*, *cementer*) appear in fewer than 5% of frames. Table I reports class-level statistics.

TABLE I: Class-level statistics for the seven classes in *Camera #2*. Left: Interventions 1–8 (training/validation). Right: Interventions 9–10 (test). % indicates the percentage of frames in which each class appears.

Class	Int. 1–8		Int. 9–10	
	# Frames	% Frames	# Frames	% Frames
instrument table	5421	100.0%	1311	100.0%
secondary table	5420	99.9%	1309	99.8%
drape	3744	69.0%	828	63.2%
hammer	221	4.1%	47	3.6%
saw	257	4.7%	66	5.0%
drill	145	2.7%	23	1.8%
cementer	115	2.1%	28	2.1%

Model Training and Inference Setup

The dataset is split at the intervention level to prevent frame overlap across training, validation, and test sets. Interventions 1–7 are used for training (4773 frames, ∼71%), Intervention 8 for validation (650 frames, ∼10%), and Interventions 9–10 for testing (1311 frames, ∼19%). We fine-tune YOLO26s-seg [1], the small variant of the YOLO26 instance segmentation architecture, initialized from COCO-pretrained weights. Training is performed for 30 epochs using an input resolution of 640×640, batch size 8, and the AdamW optimizer with weight decay of 5×10^{-4} and a 5-epoch warmup. To address class imbalance, we use a class-weighted binary cross-entropy classification loss, in which each class weight is proportional to the square root of the inverse of its training-set frequency and normalized to unit mean. At inference, YOLO26s-seg

predicts object classes, bounding boxes, and segmentation masks. For the rare classes *hammer*, *saw*, *drill*, and *cementer*, the lightweight ViT-Tiny variant of EfficientSAM [2] is selectively applied using YOLO-detected bounding boxes as spatial prompts to generate refined segmentation masks, while YOLO masks are retained for the other classes to limit additional inference cost. Experiments are conducted on an NVIDIA Tesla T4 GPU (16 GB) using the Ultralytics framework [4].

Model Evaluation

Model performance is evaluated on the held-out test set. Instance segmentation performance is measured using mask-level Average Precision (AP) following the COCO evaluation protocol [5], computed against the ground-truth annotations. Results are reported per class as mask-level AP50–95, corresponding to the average performance over Intersection over Union thresholds from 0.50 to 0.95. Inference efficiency is assessed through per-frame latency after GPU warm-up, averaged over all test frames. To evaluate the contribution of the proposed EfficientSAM refinement module, an ablation study compares the YOLO26s-seg baseline with the proposed YOLO26s-seg + EfficientSAM hybrid pipeline.

RESULTS

Per-class segmentation results on the held-out test set (Interventions 9–10) are reported in Table II.

TABLE II: Per-class mask segmentation performance on the test set. EfficientSAM refinement is selectively applied only to the selected surgical instrument classes (*hammer*, *saw*, *drill*, and *cementer*). Δ indicates the improvement obtained with EfficientSAM refinement.

Class	**YOLO AP50–95**	**YOLO+EffSAM AP50–95**	**Δ AP50–95**
instrument table	0.363	0.363	–
secondary table	0.505	0.505	–
drape	0.672	0.672	–
hammer	0.091	0.159	+0.068
saw	0.159	0.233	+0.074
drill	0.113	0.243	+0.130
cementer	0.163	0.181	+0.017

EfficientSAM improves segmentation for all four rare instrument classes, with the largest gain observed for *drill* (+0.130 AP50–95). Qualitative examples are shown in Fig. 1. The proposed YOLO26s-seg + EfficientSAM pipeline achieves a mean inference latency of 59.4 ms per frame (16.8 FPS), compared to 39.0 ms per frame (25.6 FPS) for the YOLO26s-seg baseline.

DISCUSSION AND CONCLUSION

This work presents a low-latency hybrid approach for instance segmentation of surgical tools in the OR, combining YOLO26s-seg with selective EfficientSAM refinement on the 4D-OR dataset [3]. The proposed pipeline improves mask AP50–95 for all rare instrument classes, with gains of +0.068 for *hammer*, +0.074 for *saw*, +0.130 for *drill*, and +0.017 for *cementer*. Despite these improvements, segmentation remains more accurate for large and frequently occurring scene elements than for small and underrepresented surgical instruments. A class-weighted binary cross-entropy loss is used to mitigate class imbalance during training; EfficientSAM further improves rare-class segmentation using YOLO-predicted bounding boxes as prompts. However, detector performance remains a limiting factor, since EfficientSAM can only generate masks for objects successfully detected by YOLO. Although inference speed decreases from 25.6 FPS for standalone YOLO26s-seg to 16.8 FPS for YOLO26s-seg + EfficientSAM, selective refinement restricts the additional computation to rare-class detections. Future work will investigate improved detection of challenging instruments and validation across multiple camera viewpoints. Overall, the proposed framework improves segmentation of underrepresented surgical tools while preserving low-latency inference performance, providing a low-level OR perception component for downstream tasks such as hand–instrument interaction analysis, workflow monitoring, and context-aware assistance.

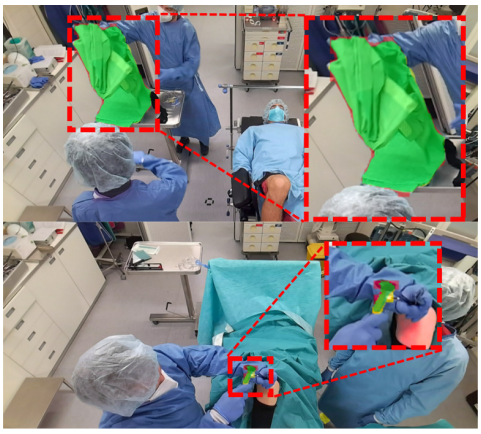


Fig. 1: Qualitative examples from the 4D-OR test set. Top: segmentation of a large scene element (*drape*). Bottom: a challenging case involving a small and rare surgical tool (*hammer*) partially occluded by clinician hands. Green indicates overlap between prediction and ground truth, yellow missed ground-truth pixels (false negatives), and red pixels incorrectly segmented as belonging to the target object (false positives).

ACKNOWLEDGMENTS

This work was supported by Fondazione Alfieri per il Cuore.

REFERENCES

[1] R. Sapkota, R. H. Cheppally, A. Sharda, and M. Karkee, "YOLO26: key architectural enhancements and performance benchmarking for real-time object detection," *arXiv preprint arXiv:2509.25164*, 2025.

[2] Y. Xiong, B. Varadarajan, L. Wu, X. Xiang, F. Xiao, C. Zhu, X. Dai, D. Wang, F. Sun, F. Iandola *et al.*, "EfficientSAM: Leveraged masked image pretraining for efficient segment anything," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, pp. 16 111–16 121.

[3] E. Özsoy, E. P. Örnek, U. Eck, T. Czempiel, F. Tombari, and N. Navab, "4D-OR: Semantic scene graphs for or domain modeling," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2022, pp. 475–485.

[4] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics YOLO," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

[5] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," *European Conference on Computer Vision (ECCV)*, pp. 740–755, 2014.

GelFlip: A Flip-Open Visual-Tactile Palpation Probe for Dual-Mode Sensing in Minimally Invasive Surgery

Chuang Lu[1], Ziting Liang[1], Tianchi Chen[1], Yuyu Cai[1], Danail Stoyanov[1], Agostino Stilli[1]

[1]*The Hawkes Institute, University College London, London, UK*
chuang.lu.22@ucl.ac.uk

INTRODUCTION

Tactile perception is essential in clinical practice: clinicians use palpation to assess tissue stiffness and texture and to detect abnormalities, including tumours. In minimally invasive surgery, however, the surgeon cannot directly palpate the operative site, so tissue condition must be inferred from limited visual cues and instrument–tissue interactions.

Previous work has explored tool-mounted force sensors, tactile arrays, and haptic interfaces to restore palpation-like feedback [1]. However, the small sensing area at an instrument tip limits spatial resolution and sensitivity. Vision-based tactile sensors such as GelSight [2] use a camera to reconstruct contact-induced surface deformation in 3-D, converting image pixels into dense sensing elements. Our previous GelPoLight sensor [3] replaces directional illumination with compact point-LED lighting for a smaller form factor and greater geometric flexibility.

Despite this progress, integrating unobstructed visual inspection and tactile sensing within a single compact instrument remains challenging. Opaque tactile surfaces may occlude the camera view, while multimodal designs may compromise one or both sensing modalities. We therefore propose GelFlip, a flip-open visual–tactile palpation probe that integrates visual inspection and tactile palpation within a single instrument for minimally invasive surgery.

MATERIALS AND METHODS

A. System Overview

GelFlip operates in two modes using the same camera. In visual mode, the tactile head flips aside for an unobstructed external view. In tactile mode, it closes and locks over the camera to capture contact-induced deformation of the reflective membrane. This enables visual inspection and tactile sensing without increasing the probe diameter.

As shown in Fig. 1, GelFlip is a cylindrical tendon-driven probe with a 25 mm outer diameter and a 300 mm overall length, comprising a 200 mm insertion section and a 100 mm handle. The handle carries optical tracking markers and houses the actuation unit. The distal tip integrates a miniature camera (XIMEA xiMU) with

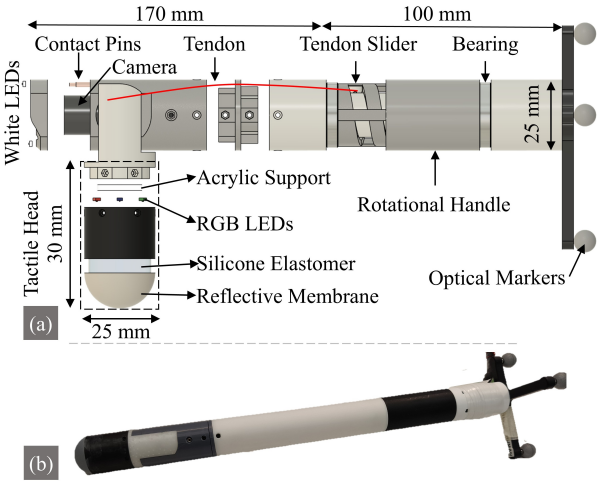


Fig. 1 GelFlip design and prototype. (a) Annotated exploded view with key dimensions (not to scale). (b) Prototype photograph.

white LEDs for endoscopic imaging. A foldable tactile head rotates within the cylindrical envelope, enabling reconfiguration between the open and closed modes without increasing the probe diameter.

B. Tactile Head

The tactile head consists of a transparent hemispherical silicone elastomer spray-coated with a thin reflective membrane to improve coating uniformity and sensitivity. Three 0201 SMD LEDs (red, green, and blue; LITE-ON Technology, Taiwan) are mounted beneath the elastomer in an equilateral triangular arrangement. Their compact footprint and wide emission angle provide tri-colour illumination across the hemispherical membrane. A transparent acrylic support maintains mechanical and optical alignment, positioning the camera beneath the elastomer to capture contact-induced deformation images.

C. Tendon-driven Flip-open Mechanism

The handle actuates the distal tactile head through two symmetric groove-slider-tendon modules. Two mirrored helical guide grooves convert handle rotation into linear slider motion. Each tendon is anchored to a slider, so the resulting tendon stroke controls the opening angle.

Integrated tensioning features allow preload adjustment to reduce slack and improve repeatability.
The two tendon paths act antagonistically on the distal pivot: rotating the handle tensions one tendon while slackening the other, thereby controlling opening and closing. The tendons attach to an offset feature on the rotating head to generate torque about the hinge axis. A quarter-circle profile with a flat face provides a hard stop at full closure and prevents over-rotation. Flat dwell segments at the groove ends lock the open and closed positions against back-driving and minimise backlash.

D. Dual-mode Visual–Tactile Registration

In visual mode, the shared camera provides an unobstructed view, while an external optical tracking system (NDI Polaris) tracks the probe's 6-DoF pose using rear-mounted markers. Multi-view images are reconstructed into a 3-D point cloud using VGGT [4]. A Sim(3) transformation estimated from the tracked poses aligns the reconstruction with the Polaris frame and recovers metric scale; sparse correspondences and robust optimisation can further reduce drift. The registered map supports target selection and tracked navigation by transforming target coordinates into the probe-tip frame. In tactile mode, GelFlip is navigated to a selected target under real-time tracking feedback. Upon contact, the reflective-membrane-coated elastomer deforms under RGB point-LED illumination and is continuously imaged by the shared camera. Per-pixel surface normals and depth can be reconstructed using the calibrated point-light photometric stereo model [3], enabling dense deformation mapping. Because the probe pose is tracked in the same Polaris frame, the reconstructed deformation can be registered to the visual map, allowing tactile findings to be localised within the visual scene.

RESULTS

A proof-of-concept study evaluated GelFlip on a liver-shaped soft phantom (Ecoflex 00-10) containing three rigid spherical inclusions, 3, 4, and 5 mm in diameter, embedded approximately 3 mm below the surface. Fiducial markers above the inclusions provided visual target cues. Fig. 2 summarises the dual-mode workflow: (a) the tactile head is opened for multi-view imaging; (b) it is closed and locked for tactile palpation; and (c) reconstruction-guided palpation is performed at visually specified targets. Under tracking feedback, the probe was navigated to each marker-defined site and palpated at the corresponding surface location. All three tactile images showed localised deformation consistent with the underlying rigid inclusions, supporting the feasibility of reaching visual targets and detecting subsurface structures not evident from surface appearance.

DISCUSSION

GelFlip integrates endoscopic inspection and targeted tactile palpation within a single flip-open instrument. The phantom study demonstrates tracked navigation to visual targets and detection of subsurface inclusions not visible endoscopically. As an early-stage prototype, its 25 mm form factor, stability during repeated switching and contact, and reliance on external optical tracking currently constrain practical use. Clinical translation will require sterilisation-compatible, biocompatible construction, reliable one-handed operation, and compatibility with minimally invasive workflows.
Future development will focus on operational robustness. Repeated-cycle and port-constrained tests will assess deployment consistency and durability, while moving, deformable phantoms will examine sensitivity to tissue motion and contact variability. Improved tracking and registration will also support reliable and repeatable target revisitation.

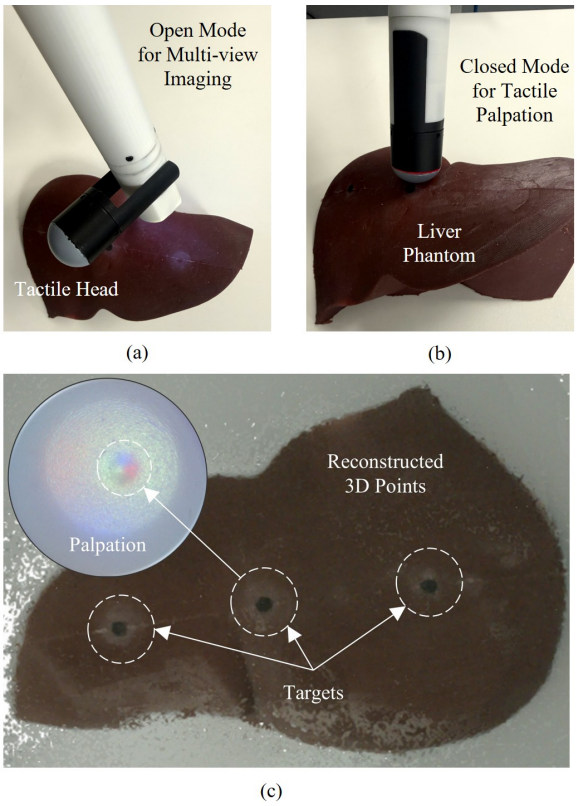


Fig. 2 Dual-mode operation and palpation workflow. (a) Open mode for multi-view imaging. (b) Closed mode for tactile palpation. (c) Reconstructed point-cloud–guided palpation at visually specified targets.

REFERENCES

[1] N. Bandari, J. Dargahi, and M. Packirisamy, "Tactile sensors for minimally invasive surgery: A review of the state-of-the-art, applications, and perspectives," *IEEE Access*, vol. 8, pp. 7682–7708, 2020.

[2] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[3] C. Lu, Z. Liang, D. Stoyanov, and A. Stilli, "Gelpolight: A novel visual tactile sensor based on photometric stereo with point lighting," *IEEE Sensors Journal*, 2025.

[4] J. Wang, M. Chen, N. Karaev, A. Vedaldi, C. Rupprecht, and D. Novotny, "Vggt: Visual geometry grounded transformer," in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 5294–5306.

Human-Robot Shared Control for Markerless Pedicle Subtraction Osteotomy Procedure

Ingrid Tombini[1,3], Lorenzo Casadio[1,3], Junling Fu[1], Francesco Costa[2], and Elena De Momi[1]

[1]*Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy*
[2]*IRCCS Fondazione Istituto Neurologico Carlo Besta, Milano, Italy*
[3]*These authors contributed equally to this work*

INTRODUCTION

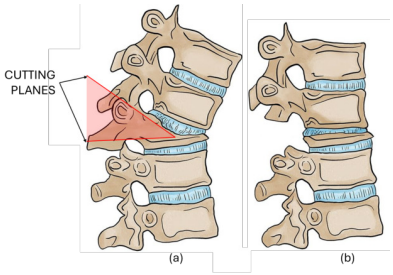


Fig. 1: Simplified representation of PSO. (a) Resection planes. (b) Closure of the osteotomy and restoration of alignment.

Spinal osteotomy involves controlled resection of vertebral bone to correct pathological spinal curvature. Among available techniques, pedicle subtraction osteotomy (PSO) is the procedure of choice for severe, rigid sagittal imbalance [1]. In PSO, a triangular wedge of bone is resected through the pedicles and vertebral body, and the gap is closed under compression to restore sagittal alignment (Fig. 1). Pedicle screws inserted above and below the osteotomy provide mechanical stability and enable controlled closure through rod compression. Despite its effectiveness, PSO remains highly demanding. Neurological complications have been reported in up to 11% of cases [2], and progressive anatomical changes during resection and closure make intraoperative spatial awareness particularly challenging.

Robotic shared-control frameworks have demonstrated improvements in cutting accuracy for osteotomy procedures [3], commonly relying on optical tracking with reference frames attached to the patient [4]. However, progressive bone resection and osteotomy closure make such references difficult to maintain, and line-of-sight interruptions further compromise reliability [4].

We propose a shared control system for PSO that avoids dedicated optical fiducials by tracking pedicle screw heads (surgical hardware already implanted for fixation) as intrinsic landmarks, enabling continuous registration update without markers attached solely for tracking purposes.

MATERIALS AND METHODS

The proposed system integrates a preoperative planning stage with a physical robotic platform, as illustrated in Fig. 2.

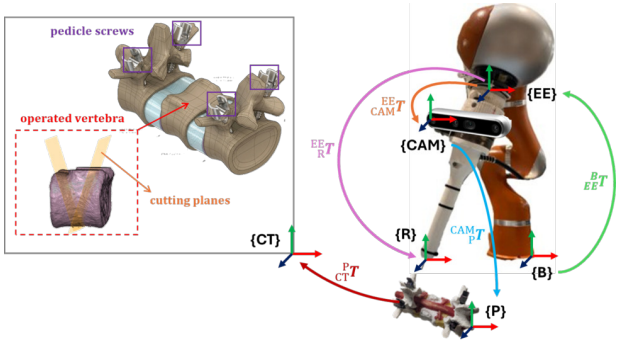


Fig. 2: Overview of the proposed system. Left: preoperative plan $\{CT\}$ showing PSO cutting planes. Right: robotic platform with the RGB-D camera $\{CAM\}$ and drill $\{R\}$ rigidly mounted on the end-effector $\{EE\}$, and the 3D-printed phantom $\{P\}$ at the base $\{B\}$.

A. Preoperative Planning

As depicted in Fig. 2 left, a spinal segment is reconstructed from CT data using 3D Slicer. The CT dataset includes pedicle screws, highlighted in the figure, whose head centroids are annotated to serve as fiducial landmarks. Based on the selected osteotomy level, a wedge resection is planned by defining two cutting planes, shown in orange in Fig. 2, that determine the desired correction angle. This establishes a known geometric correspondence between screw centroids and osteotomy planes in the CT reference frame, which serves as the preoperative plan.

B. Screw Detection and Tracking

Intraoperative screw detection is performed using a YOLOv8 model trained on 1500 annotated RGB images of a phantom ($\{P\}$ in Fig. 2). The network detects a single class ("Screw") without enforcing any predefined ordering, as all screws are visually indistinguishable. Screw identities are initialised by ordering detections geometrically in the camera frame, assuming a fixed patient-camera orientation, and subsequently maintained via optical flow tracking.

A hand-eye calibration estimates ${}^{EE}_{CAM}\mathbf{T}$, the rigid transformation between camera and end-effector. Combined with the forward kinematics ${}^{B}_{EE}\mathbf{T}$, each screw position ${}^{\mathrm{CAM}}\mathbf{p}_i$ is expressed in the robot base frame $\{\mathrm{B}\}$:

$$ {}^{B}\mathbf{p}_i = {}^{B}_{EE}\mathbf{T} \cdot {}^{EE}_{CAM}\mathbf{T} \cdot {}^{CAM}\mathbf{p}_i \tag{1} $$

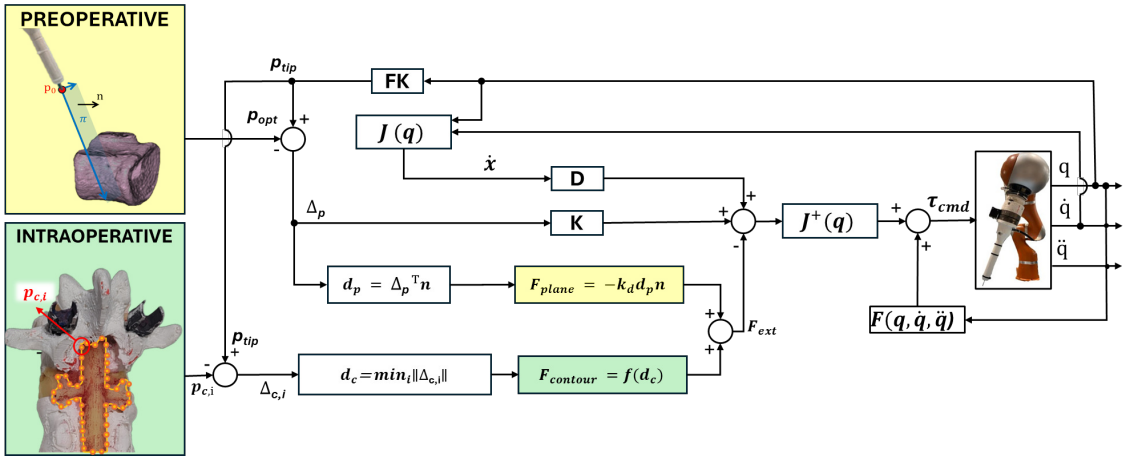


Fig. 3: Shared control scheme. Preoperatively, the cutting plane is defined by entry point $\mathbf{p}_0$ and normal $\mathbf{n}$. Intraoperatively, the drill tip position $\mathbf{p}_{tip}$ and velocity $\dot{\mathbf{x}}$ are obtained via forward kinematics and Jacobian $\mathbf{J}(\mathbf{q})$. Two virtual fixture branches compute: (i) $\mathbf{F}_{plane} = -k_d d_p \mathbf{n}$, attracting the tool to the cutting plane, and (ii) $\mathbf{F}_{contour} = f(d_c)$, a sigmoidal repulsive force preventing drill encroachment on the segmented spinal cord contour. Forces are combined with damping $\mathbf{D}$ and stiffness $\mathbf{K}$, then mapped to joint torques via $\mathbf{J}^+(\mathbf{q})$.

Finally, intraoperative registration is performed by computing the rigid transformation ${}^{P}_{CT}\mathbf{T}$ that aligns the preoperative screw centroids with the detected ones using Singular Value Decomposition (SVD).

C. Shared Control and Safety Constraints

The control operates in two sequential phases. Upon pedal activation, the robot autonomously reaches the entry point $\mathbf{p}_0$ via a fifth-order polynomial trajectory. After convergence, the controller switches to shared-control mode based on Cartesian impedance with anisotropic stiffness.

The cutting plane is defined by entry point $\mathbf{p}_0 \in \mathbb{R}^3$ and unit normal $\mathbf{n} \in \mathbb{R}^3$. An attractive virtual fixture constrains the tool to the plane via:

$$\mathbf{F}_{plane} = -k_d \, d_p \, \mathbf{n}, \quad d_p = \Delta_p^{\top}\mathbf{n} \tag{2}$$

where $k_d > 0$ is a stiffness gain and $\Delta_p = \mathbf{p}_{tip} - \mathbf{p}_{opt}$, with $\mathbf{p}_{opt}$ any point on the cutting plane and $\mathbf{p}_{tip}$ the current drill tip position.

A YOLOv8 segmentation model identifies the spinal cord contour $\{\mathbf{p}_{c,i}\}$ and projects it into the robot base frame as a dynamic no-fly zone. Defining $d_c = \min_i \|\mathbf{p}_{tip} - \mathbf{p}_{c,i}\|$ as the distance to the nearest contour point, a repulsive virtual fixture is activated via $\mathbf{F}_{contour} = f(d_c)$ generating increasing lateral repulsion as the drill approaches the spinal cord boundary, while remaining constrained to the osteotomy plane.

EXPERIMENTAL PROTOCOL AND RESULTS

A PLA phantom was 3D-printed from CT data, with silicone intervertebral discs and spinal cord, and pedicle screws manually inserted per the preoperative plan. Both models were trained on 1500 annotated RGB images (100 epochs, Adam, lr 10^{-3}).

Hand-eye calibration. Validation yielded a reprojection error of 0.24 px and a mean Euclidean error of 1.21 mm in the robot base frame.

Screw detection and spinal cord segmentation. The detection model achieved high accuracy (precision 0.992, recall 0.964, mAP@50 0.994). Segmentation performance was strong at loose IoU thresholds (mAP@50 0.895) but degraded at stricter ones (mAP@50–95 = 0.178), limiting precise safety boundary definition.

Intraoperative registration. Alignment of preoperative and intraoperative screw centroids yielded RMSE = 1.23 ± 0.24 mm, consistent with manual screw insertion accuracy.

Shared control. Ten non-expert users performed a simulated osteotomy. Mean plane deviation was 0.61 ± 0.46 mm, within clinically acceptable limits and well below freehand errors of 2.27-4.53 mm reported for a comparable setup [3]. The no-fly zone was violated in one trial; otherwise a minimum distance of 0.20 ± 0.11 mm from the spinal cord was maintained.

CONCLUSIONS

A markerless shared control framework for PSO was validated, achieving sub-millimetric cutting precision and accurate registration. Segmentation accuracy remains the main limitation for reliable safety boundary definition. Future work will target improved segmentation on real intraoperative data and more realistic phantom development.

REFERENCES

[1] K. C. Kose, O. Bozduman, A. E. Yenigul, and S. Igrek, "Spinal osteotomies: indications, limits and pitfalls," *EFORT open reviews*, vol. 2, no. 3, pp. 73–82, 2017.

[2] J. M. Buchowski, K. H. Bridwell, L. G. Lenke, and et al., "Neurologic complications of lumbar pedicle subtraction osteotomy: a 10-year assessment," *Spine (Phila Pa 1976)*, vol. 32, no. 20, pp. 2245–2252, 2007.

[3] E. Iovene, L. Casadio, J. Fu, F. Costa, G. Ferrigno, and E. De Momi, "Human–robot shared control for osteotomy procedure," *IEEE Transactions on Biomedical Engineering*, vol. 72, no. 9, pp. 2698–2708, 2025.

[4] F. Manni, A. Elmi-Terander, G. Burström, O. Persson, E. Edström, R. Holthuizen, C. Shan, S. Zinger, F. van der Sommen, and P. H. N. de With, "Towards optical imaging for spine tracking without markers in navigated spine surgery," *Sensors*, vol. 20, no. 13, p. 3641, 2020.

Electrical Impedance Tomography in Constrained Endoscopy Applications

Chiara Falagario[1,2], Zhuoqi Cheng[3], Elena De Momi[2], Leonardo S. Mattos[1]

[1]*Biomedical Robotics Lab, Istituto Italiano di Tecnologia, Genoa, Italy, chiara.falagario@iit.it*

[2]*Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy*

3Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Odense, Denmark

INTRODUCTION

Head and neck squamous cell carcinomas is the sixth most common cancer globally, with incidence projected to increase by approximately 30% by 2030 [1]. Tumours of the pharyngeal and laryngeal regions are particularly challenging because they can only be evaluated through laryngoscopic examination [2]. Still, accurate tumoral margin delineation is essential to ensure complete tumour removal while preserving healthy tissue.

Several techniques have been investigated for tumour detection [3] and surgical margin assessment [4], among which Electrical Impedance Tomography (EIT) has recently emerged as a promising label-free imaging modality for different diagnostic tasks [5,6]. However, existing EIT studies have not addressed tumours of the pharyngeal and laryngeal regions, where access is restricted to flexible endoscopy. The severe constraints imposed by the working channel of the tool drastically limit the number of electrodes that can be integrated on the probe tip, reducing the spatial information available for conventional EIT reconstruction.

In this context, the objective of this work is to design a reliable EIT-based tumour reconstruction method specific for constrained endoscopic applications. The proposed approach integrates measurements acquired across multiple positions on the tissue surface, leveraging probe motion to synthesize a larger virtual sensing aperture despite the limited number of electrodes.

MATERIALS AND METHODS

A micro-laryngoscopic probe based on the SmartProbe geometry [7] was modelled as a planar array (dimensions compatible with the working channel of flexible laryngoscopes) of four electrodes placed to increase the difference between sensitivity distributions for different measurements on the surface of the test volume.

The tissue volume under test was modelled as a 3D homogeneous cuboid (σ = 0.05 S/m) containing a centrally located cylinder, representing the simulated tumoral mass (STM), with a clinically relevant conductivity contrast [8] with respect to the background (σ = 2.2 S/m).

Two experiments were conducted: one in simulation and one using an experimental phantom. The phantom (Figure 1a) was realized assembling two agar gels with different saline concentrations, to mimic mucosal tissue with a tumoral mass (Figure 1b). In both experiments, the workflow began with the acquisition of bioimpedance signals at 9 positions, defined on a 3×3 grid on the model surface, for three times. Each position was defined by a contact centre and the four theoretical relative positions of the electrodes constituting the probe.

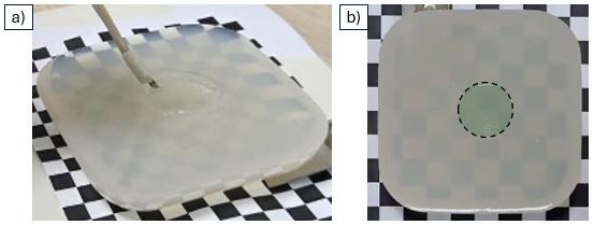


Figure 1. Agar phantom used in the experiments. a) Side view of the phantom with contacting probe. b) Top view of the phantom with the STM highlighted in a dashed green circle.

During the experiments with the agar phantom, probe positioning has been performed using a robotic micromanipulator (Siskiyou MX7600L) for controlled placement. Probe positions were inferred using a marker-less YOLOv11-based key point detection approach, which is potentially also applicable to the envisioned clinical scenario. This enabled the localization of the tip centre and computation of its distance from the reference origin in the centre of the gel phantom. Figure 2a presents the experimental setup for data acquisition and probe localization, which was made up of the robotic micromanipulator controlled by a PC, a camera, and the SmartProbe Processing Unit (a touch screen computer and a bioimpedance spectroscopy sensor). Figure 2b depicts the grid of acquisition poses on the agar phantom used for EIT reconstruction, together with the theoretical electrode layout for one representative probe position.

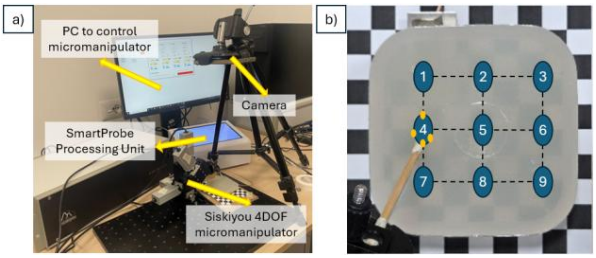


Figure 2. Experiments with the agar phantom. a) Complete experimental setup. b) Acquisition grid defined for the reconstruction (blue) with electrode layout model (yellow).

The experiments in the simulated environment were performed using the same positions acquired by the marker-less approach to model the electrodes' layout. To simulate manufacturing tolerances and positioning uncertainty, controlled noise was introduced into the simulation to enforce a prescribed signal-to-noise ratio.

For each probe pose, the proposed method adopted a monopolar stimulation scheme, which considered one

active electrode at a time referenced to the ground. Once all the acquisitions were completed, the signals were arranged into a unified measurement array to emulate a larger probe configuration. Measurements from a single position were replicated to each position of the acquisition grid, thereby defining a spatially homogeneous reference state used to perform differential reconstruction. Finally, the conductivity distribution was reconstructed using a coarse-to-fine approach and the results were visualized on a superficial 2D slice extracted from the mesh.

To evaluate the performance of the proposed method, overlap-based and localization-based metrics were considered. The Dice Similarity Coefficient $DSC = \frac{2|A \cap B|}{|A|+|B|}$ was used to assess the overlap between reconstructed (A) and ground truth (B) STM regions.

Localization accuracy was evaluated through the Position Error $PE = \left\| c_{rec} - c_{gt} \right\|_2$, defined as the Euclidean distance in millimetre between the centroids of the reconstructed (c_{rec}) and ground truth (c_{gt}) STM.

Two segmentation procedures were adopted to generate the STM masks to compute the metrics. The ground truth mask was manually defined on the model surface, whereas the masks for reconstructed slices were obtained through an automated segmentation pipeline involving relative intensity thresholding and extraction of the connected components containing the global maximum.

RESULTS

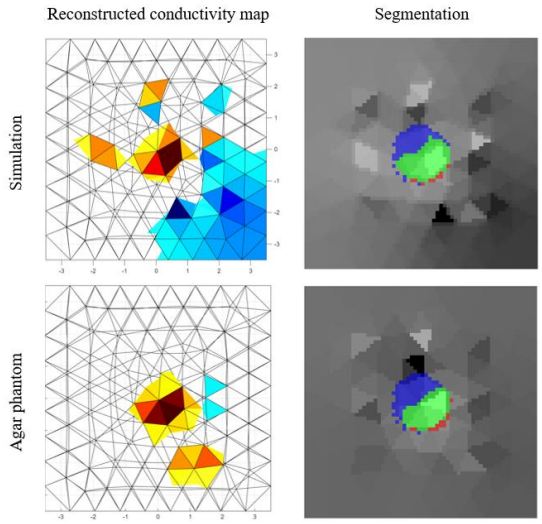


Figure 3. Results for the experiments in simulation (top) and on the agar phantom (bottom). Left: reconstructed conductivity distribution maps. Right: segmentation masks showing ground truth (blue), reconstructed STM (red) and their overlap (green).

Figure 3 shows the results obtained during the experiments. In simulation (upper row), the reconstructed conductivity map exhibited a well-localized STM, yielding a DSC = 0.660 and a PE = 2.338 mm. In the agar phantom experiments using SmartProbe (lower row), across the 3 repeated acquisitions, the method achieved a mean DSC = 0.725 ± 0.023 and a mean PE = 1.990 ± 0.176 mm, indicating consistent overlap with the ground truth and accurate localization of the STM.

CONCLUSION AND DISCUSSION

Given the constraints on probe size and electrode number imposed by flexible endoscopes for laryngoscopy, the EIT reconstruction method proposed in this work shows strong potential for delineating tumoral margins by leveraging probe motion to enhance spatial information. Combined analysis of overlap and localization demonstrates that the method achieves consistent and reliable results, comparable to those reported in the literature, in identifying the simulated tumoral mass and estimating its position with good spatial agreement and limited localization error.

Residual discrepancies in boundary delineation could be mainly attributed to measurement noise and a limited number of probe positions. Improving the reconstruction accuracy will be addressed in future studies by increasing the number of positions and by integrating strategies to identify optimal probe placements that maximize the information content of the measurements.

Overall, the results highlight the potential of the proposed approach for accurate localization and delineation of tumoral regions, supporting its feasibility for clinical application even in anatomically challenging areas such as the larynx and pharynx.

ACKNOWLEDGMENTS

This work was partially funded by the European Union - Horizon Europe project “AIRCARE - AI-augmented Robotics for CAncer point of caRE” (101137426).

REFERENCES

[1] Ferlay, Jacques, et al. "Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods." *International journal of cancer* 144.8 (2019): 1941-1953.

[2] Johnson, Daniel E., et al. "Head and neck squamous cell carcinoma." *Nature reviews Disease primers* 6.1 (2020): 92.

[3] Gerstner, Andreas O. H. “Early detection in head and neck cancer - current state and future perspectives.” *GMS Curr Top Otorhinolaryngol Head Neck Surg* **7**, Doc06 (2008).

[4] Chen, Yang., et al. “Surgical margins in head and neck squamous cell carcinoma: A narrative review.” *International Journal of Surgery* 110, 3680–3700 (2024).

[5] Gatabi, Zahra Rezanejad, et al. "Application of Electrical Impedance Tomography in Diagnosis of Cancer." *Advanced Biomedical Research* 14.1 (2025): 136.

[6] Halter, Ryan J., and Young-Joong Kim. "Toward microendoscopic electrical impedance tomography for intraoperative surgical margin assessment." IEEE Transactions on Biomedical Engineering 61.11 (2014): 2779-2786.

[7] Petersen, Marianne Viola, Thiusius Rajeeth Savarimuthu, and Zhuoqi Cheng. "Development of a Bioimpedance Probe for Enhanced Tissue Identification in the Upper Aerodigestive Tract." *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2025.

[8] Hancu, Ileana, et al. "Electrical properties tomography: Available contrast and reconstruction capabilities." *Magnetic resonance in medicine* 81.2 (2019): 803-810.

Deep Learning-Based contact detection for Robot-Assisted retinal surgery using microscope and OCT imaging

Giulia Galvan[1], Ivo De Clerck[2], Joy Haegemans[1], Mouloud Ourak[1], and Emmanuel Vander Poorten[1]
[1]Department of Mechanical Engineering, KU Leuven, Leuven, Belgium
[2]Department of Ophthalmology, University Hospitals Leuven, Belgium

INTRODUCTION

Age-related Macular Degeneration (AMD) is a progressive chronic eye disease and the leading cause of severe irreversible central vision loss among older adults [1]. Depending on whether the condition is of the "dry" or "wet" type and the stage of the disease, the treatment can vary. In selected cases of wet AMD, such as subretinal hemorrhage, subretinal injection of tissue-Plasminogen Activator (tPA) represents a possible surgical approach [2]. Manually performing a subretinal injection requires micrometer-scale precision, it is highly sensitive to physiological hand tremor, and involves prolonged injection durations [3]. For this reason, robotic assistance has become a key technology for improving precision and expanding intraocular procedures, but it does not inherently enhance depth perception [4], [5]. This limitation is further compounded by the stereo-microscope, the imaging modality routinely used during eye surgery, that offers a convenient visual stream but provides limited depth perception and does not allow visualization beyond the retinal surface. On the other hand, intraoperative Optical Coherence Tomography (iOCT) imaging modality, provides high-resolution cross-sectional imaging, enabling visualization of subsurface layers. However, conventional OCT scanning is not sufficiently fast for real-time instrument tracking and constrained by a limited field of view, requiring prior knowledge of the needle position or motion to maintain reliable tracking during surgical maneuvers [6]. Deep learning (DL) has been increasingly explored to provide real-time guidance during ophthalmic surgery and prior work has demonstrated the use of ResNet-based models for instrument tracking from microscope images [7]. This has led to explore the use of similar DL approaches on iOCT data to detect needle–retina interactions [8]. Accurate detection of needle–retina contact is critical, as retinal tissue is highly fragile and even minimal unintended contact may result in damage. Building on this need and previous developments, this work represents an initial step toward a multimodal guidance pipeline that integrates both microscope and iOCT data. In this stage, the focus is on validating a ResNet-based method on microscope data and exploring its extension to needle detection in iOCT images, with particular attention to model interpretability to identify the regions driving the network's predictions.

MATERIALS AND METHODS

A ResNet18 architecture [9], selected among ResNet variants including ResNet50 and ResNet101, was employed to classify frames into 'contact' and 'no contact' states for the microscope images and 'needle' and 'no needle' for the OCT scans (pipeline showed in Fig. 1). The network was initialised with pre-trained weights on ImageNet, using transfer learning as a starting point for feature extraction. In the final configuration, the backbone parameters were frozen, while only the final convolutional block (Layer 4) and the fully connected layer were optimized using Stochastic Gradient Descent (SGD) with a learning rate of 0.001 and weight decay of 0.0005. The dataset was developed from 16 surgical videos of robot-assisted subretinal injections performed under general anesthesia using the DORC EVA Nexus platform, ZEISS Artevo microscopes, and iOCT monitoring. The videos were acquired as part of RoTIS study that followed the Declaration of Helsinki, was approved by relevant ethics committees (UZ Leuven ethics committee S67215; Federal Agency for Medicines and Health Products CIV-23-12-045354), and registered at ClinicalTrials.gov (NCT06347107). From these videos, two distinct datasets were constructed. For the first dataset, only the microscope images were considered and a total of 1,719 frames were extracted; to focus the analysis on the surgical task, a specific region of interest (ROI) centered around the needle was selected for each frame. The second dataset was derived from iOCT scans, comprising a total of 520 frames. Both the dataset were partitioned into training (70%), validation (20%), and testing (10%) sets. Explainable artificial intelligence (XAI) was further investigated as an added contribution of this work through Grad-CAM visualizations, which were employed to qualitatively assess the regions driving the network's predictions [10]. Explainability is particularly important in safety-critical surgical applications, as it enables verification that model decisions are based on clinically relevant features rather than spurious visual cues. Model performance on the independent test set was evaluated using a confusion matrix and a classification report, including precision, recall, and F1-score. Overall accuracy was also computed to summarize the results.

RESULTS

The ResNet18 model with a frozen backbone was first evaluated on microscope images using a train/validation

split, as done in literature [7], achieving an overall accuracy of 0.96, with macro-averaged precision, recall, and F1-score of 0.96, consistent with results reported. When evaluated on an independent test set, the model maintained a high performance, reaching an accuracy of 0.88, with overall precision, recall, and F1-score of 0.88. Additional architectures were also evaluated on microscope images. For example, the ResNet50 achieved an accuracy of 0.83 with an overall F1-score of 0.80, while ResNet101 obtained an accuracy of 0.81 and an F1-score of 0.77. Based on these results, the ResNet18 architecture was selected for further experiments on iOCT images. On this modality, the model achieved an accuracy of 0.94, with overall precision, recall, and F1-score of 0.94 on the test set. Grad-CAM visualizations on iOCT images show that, in the presence of the needle, the model focuses on a localized region corresponding to the needle structure. In contrast, when the needle is absent, the activation is more broadly distributed across the retinal layers. Grad-CAM visualizations were also evaluated on microscope images, where the model was observed to focus on the needle region in the majority of cases. However, in some instances, the attention was directed toward retinal vessels or hemorrhagic areas instead of the needle. Inference time was evaluated on both modalities, achieving an average latency of 6.60 ms (±0.53 ms) on OCT images, and 7.11 ms (±0.43 ms) on microscope images on the NVIDIA Clara AGX platform.

CONCLUSIONS AND DISCUSSION

The results demonstrated that Resnet18, successfully applied to microscope-based contact detection, can be effectively extended to iOCT data. The integration of both imaging streams could be particularly relevant since the microscope images can be exploited to track the needle position and dynamically reposition the iOCT scan window, while tracking the needle on the OCT scans can be useful especially in the subretinal layers. The Grad-CAM analysis provided evidence that the network relies on anatomically meaningful features to drive its predictions, and this represented an important validation step. The occasional misattribution of attention toward retinal vessels or hemorrhagic areas in microscope images underlined the importance of incorporating XAI tools to monitor the reliability of the trained models. The comparison across architectures confirmed that increasing model complexity does not translate into performance gains. Building on these findings, future work will focus on extending the pipeline toward contact detection on iOCT images, bringing the system closer to providing complete, real-time surgical guidance for robot-assisted subretinal injection.

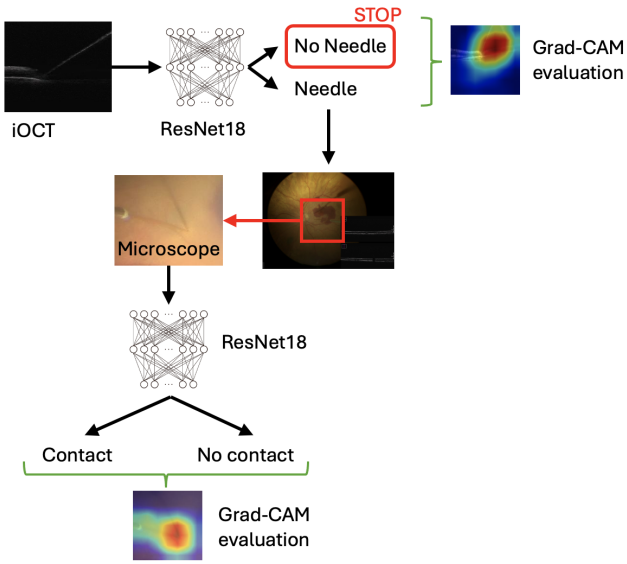


Fig. 1: Proposed pipeline for needle detection from iOCT images followed by contact detection from microscope images.

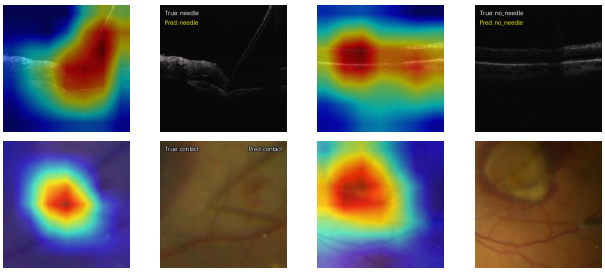


Fig. 2: Grad-CAM visualizations for OCT needle detection (top row: needle presence on the left and needle absence on the right) and microscope-based contact detection (bottom row: the model correctly focuses on the needle region on the left, the activation is mainly directed toward the hemorrhagic area, on the right). Warmer colors indicate image regions that contribute more strongly to the network prediction.

REFERENCES

[1] R. H. Guymer. Age-related macular degeneration. *The Lancet*, 2023.

[2] Novelli FJD. A new method of subretinal injection of tissue plasminogen activator and air in patients with submacular hemorrhage. *Retina*, 2017.

[3] R. Ladha. Subretinal therapy: Technological solutions to surgical and immunological challenges. *Frontiers in Medicine*, 2022.

[4] Vander Poorten Emmanuel. Robotic retinal surgery. In *Handbook of Robotic and Image-Guided Surgery*, pages 627–672. Elsevier, 2020.

[5] Ivo De Clerck. In-human feasibility and safety of subretinal drug injection through attached retina using a robotic comanipulation system. *Retina*, 2025.

[6] P. Xu. Fast oct-based needle tracking for retinal microsurgery using dynamic spiral scanning. *IEEE Transactions on Medical Robotics and Bionics*, 6(4):1502–1511, November 2024.

[7] P. F. Baldi. Vitreoretinal surgical instrument tracking in three dimensions using deep learning. *Translational Vision Science & Technology*, 2023.

[8] M. Zhou. Needle detection and localisation for robot-assisted subretinal injection using deep learning. *CAAI Transactions on Intelligent Technology*, 2025.

[9] K. He. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[10] R. R. Selvaraju. Grad-cam: Visual explanations from deep networks via gradient-based localization. *IEEE International Conference on Computer Vision (ICCV)*, 2017.

Weakly Supervised Depth Estimation and Control of a Micropipette

Huanyu Tian[1], Martin Huber[1], Lingyun Zeng[1], Zhe Han[1], Wayne Bennett[2], Giuseppe Silvestri[2], Gerardo Mendizabal-Ruiz[2], Tom Vercauteren[1], Alejandro Chavez-Badiola[3,2], and Christos Bergeles[1,2]

[1]*School of Biomedical Engineering & Imaging Sciences, King's College London, UK* (huanyu.tian@kcl.ac.uk); [2]*Conceivable Life Sciences, New York City, US and London, UK;* [3]*Hope IVF Mexico, Mexico.*

INTRODUCTION

Robotic assistance under microscopic guidance requires sub-millimetre 3D depth for cell handling tasks [1], yet depth estimation faces two fundamental obstacles: (i) transparent micropipettes violate the photo-consistency assumption of stereo methods; (ii) monocular defocus cues are weak, as depth-of-field is shallow, gradients vanish far from the focal plane [2], and symmetric blur makes directional disambiguation unreliable.

Existing methods resort to iterative focal stacking or hill-climbing sharpness searches [3], [4], which are slow and prone to local optima. Supervised alternatives [5], [6] assume synthetic Gaussian blur, which breaks down at air-liquid interfaces; annotated datasets are also scarce due to ambiguous boundaries and glare.

We address these via a *weakly supervised* strategy on the steady-hand co-manipulation paradigm [7]. Short "warm-up" trajectories under user guidance, comprising lateral FoV sweeps for hand-eye calibration and axial scans to excite defocus variation, provide kinematically grounded depth labels without manual annotation. A compact Depth Estimation Block (DEB) trained on the resulting pseudo-labeled dataset drives a confidence-gated micro-controller for real-time axial regulation.

MATERIALS AND METHODS

A. Kinematic Depth Label Generation

2D Tip positions in monocular data stream $\hat{\mathbf{p}}_{\text{tip},t}$ are obtained from a SAM2-based [8] pipeline that propagates bidirectional masks over fixed-length clips from sparse anchor-frame prompts. During a dedicated axial warm-up scan, the robot sweeps the tip along the microscope's optical (z) axis (see Fig. 1) at pixels $\{a_k\}$ named "anchors". Let $\mathcal{T}_a = \{t : \|\hat{\mathbf{p}}_{\text{tip},t} - a_k\|_2 \le \varepsilon\}$ collect frames from the dedicated sweep at anchor a_k. For each anchor, the focal plane (reference $z{=}0$) is identified as the frame of maximum image sharpness:

$$t_a^\star = \arg\max_{t\in\mathcal{T}_a} \frac{1}{|\Omega|} \sum_{(i,j)\in\Omega} \left(G_x^t(i,j)^2 + G_y^t(i,j)^2\right), \quad (1)$$

where Ω indexes an $L{\times}L$ patch around the detected tip and (G_x^t, G_y^t) are Sobel gradients. Each frame $t \in \mathcal{T}_a$ is assigned a *relative* pseudo-label $\Delta z_t := z_t - z^\star$, the signed stage displacement from the focal plane; this self-referenced strategy avoids calibrating an absolute defocus-to-metric-depth mapping. Refractive-index variation rescales this mapping but only affects convergence speed: since the objective is $\Delta z \to 0$, any $\alpha > 0$ satisfies $\alpha\Delta z \to 0 \Leftrightarrow \Delta z \to 0$ and is absorbed into K_z.

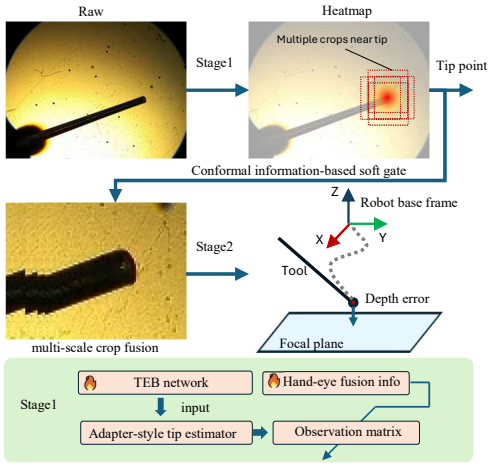


Fig. 1: Overview of the proposed perception and control pipeline. Tip esimation block (TEB) refers to the tip detection task.

B. Depth Estimation Block (DEB)

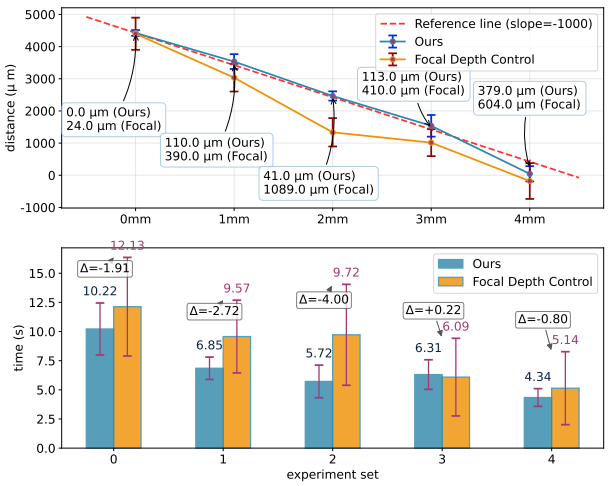


Fig. 2: Depth regulation performance (7 trials each). *Top:* Final positioning error. *Bottom:* Convergence time. Our DEB-based method (blue) shows a quasi-linear descent and converges faster, whereas the iterative baseline (orange) exhibits non-linear degradation and higher latency.

A compact ResNet18 network processes 80×80 px grayscale crops centered at the detected tip, projecting multi-scale pooled features to an L2-normalized descriptor $\boldsymbol{\psi} \in \mathbb{R}^d$ via an MLP embedding head. Two lightweight heads decode $\boldsymbol{\psi}$: a depth regression head $\hat{z} = f_d(\boldsymbol{\psi}) \in \mathbb{R}$ and an auxiliary classification head

$$\hat{\boldsymbol{\zeta}} = \mathrm{softmax}\big(f_c(\boldsymbol{\psi})\big) \in \mathbb{R}^3, \tag{2}$$

predicting `below`/`near`/`above` the focal plane (`near`: within the $76\,\mu$m depth of field of the $2\times$ objective) as a coarse gain-scheduling signal for the controller.

Training minimizes $\mathcal{L}_{\mathrm{mix}}$ (Eq. 3), where $\mathcal{L}_{\mathrm{clean}}$ encapsulates (i) depth regression (MSE), (ii) an embedding-depth k-NN alignment term [9], and (iii) classification cross-entropy; the adversarial term $\lambda^z_{\mathrm{cons}}\|\hat{z}(P_{\mathrm{adv}}) - \hat{z}(P)\|_1$ penalizes prediction changes under PGD [10] perturbations, stabilizing depth-MAE under limited warm-up data:

$$\begin{aligned}\mathcal{L}_{\mathrm{mix}} = (1-\gamma)\,\mathcal{L}_{\mathrm{clean}}(P) + \gamma\,\mathcal{L}_{\mathrm{clean}}(P_{\mathrm{adv}}) \\ + \lambda^z_{\mathrm{cons}}\,\big\|\hat{z}(P_{\mathrm{adv}}) - \hat{z}(P)\big\|_1.\end{aligned} \tag{3}$$

C. Depth-Dominant Micro-Controller

The DEB output feeds a dual-branch Kalman filter tracking planar $\mathbf{e}^{(x)}_t \in \mathbb{R}^4$ and axial $\mathbf{e}^{(z)}_t = [e_{z,t}, \dot{e}_{z,t}]^\top$ error states which serves as a command value for depth controller; up to five spatially offset crops are averaged per step to reduce single-crop estimation noise.

Rapid lateral motion induces motion blur that destabilizes the DEB; to decouple the axes, the controller switches between two modes:

- *Lateral-Dominant*: while converging laterally, axial depth is locked at the reference by a virtual fixture

$$u^z_t = \mathrm{sat}_{v^z_{\max}}\big(K_z\,\mathbf{e}^{(z)}_{\mathrm{reg},t}\big), \tag{4}$$

  where $\mathbf{e}^{(z)}_{\mathrm{reg},t}$ is the plane-keeping error (positional drift from the depth at motion onset, augmented by axial velocity) and $\mathrm{sat}_{v^z_{\max}}$ saturates velocity to prevent overshooting.
- *Depth-Dominant*: once the lateral error $\|\hat{\mathbf{e}}^{(x)}_t\|$ falls within a dead-zone tolerance, lateral motion is suppressed and the live DEB estimate drives axial regulation

$$\mathbf{u}^x_t = \mathbf{0}, \tag{5}$$

$$u^z_t = \mathrm{sat}_{v^z_{\max}}\big(K_z\,\hat{e}^{(z)}_t\big). \tag{6}$$

The micro velocity $\mathbf{u}^{\mathrm{micro}}_t = [\mathbf{u}^x_t,\, u^z_t]^\top$ is gated by $\mathcal{I}_{\mathrm{vis}}$ and superimposed onto the admittance macro command via a damped pseudo-inverse Jacobian.

CONCLUSIONS AND DISCUSSION

We evaluated the depth controller against an iterative focal-depth baseline using a motorized linear stage imposing displacements of 0–4 mm (7 repeated trials per displacement with randomized lateral starts). As shown in Fig. 2, our method achieves $41\,\mu$m vs. $1089\,\mu$m at 2 mm and $379\,\mu$m vs. $604\,\mu$m at 4 mm, with faster convergence (5.72 s vs. 9.72 s) since the DEB directly regresses depth rather than performing an iterative search.

In 3D closed-loop trials combining lateral servoing and axial regulation, the worst-case 95% confidence bound on absolute depth error is $\mathbf{291}\,\mu$m, sufficient for cell-transportation tasks: although biological targets (oocytes) measure 100–150 μm in diameter, the residual axial error is mitigated by hydraulic aspiration pressure at the pipette tip [1].

The proposed approach eliminates manual depth annotation and external sensors; by anchoring labels to the self-identified focal plane, it adapts to any optics combination through a one-time ($\sim$22 min) warm-up, making it deployable across laboratory setups.

REFERENCES

[1] A. Chavez-Badiola, G. Mendizabal-Ruiz, A. Flores-Saiffe Farías, N. Costa-Borges, A. Murray, M. Alikani, G. Silvestri, C. Millan, E. Hernández-Morales, R. Valencia-Murillo, *et al.*, "Automated oocyte retrieval, denudation, sperm preparation, and icsi in the ivf laboratory: a proof-of-concept study and report of the first live births," *Human Reproduction*, p. deaf240, 2025.

[2] S. Gur and L. Wolf, "Single image depth estimation trained via depth from defocus cues," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 7683–7692.

[3] Y. Zhang, A. Xu, Y. Chen, M. Q.-H. Meng, and L. Liu, "Multi-scale visual servoing framework for optical microscopy based on sift matching," *IEEE Robotics and Automation Letters*, vol. 8, no. 12, pp. 8144–8151, 2023.

[4] S. Pertuz, D. Puig, and M. A. Garcia, "Analysis of focus measure operators for shape-from-focus," *Pattern Recognition*, vol. 46, no. 5, pp. 1415–1432, 2013.

[5] Y. Luan, Y. Luo, Y. Liu, M. Zhang, Y. An, J. Yang, Y. Guo, and G.-Z. Yang, "Autofocusing with 3-d tracking for robot-assisted microsurgery," *IEEE/ASME Transactions on Mechatronics*, 2025.

[6] H. Yang, Y. Jin, G. Shan, Y. Wang, Y. Zheng, J. Yu, Y. Sun, and Z. Zhang, "Weakly-supervised depth completion during robotic micromanipulation from a monocular microscopic image," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 15 615–15 621.

[7] A. Bettini, P. Marayong, S. Lang, A. Okamura, and G. Hager, "Vision-assisted control for manipulation using virtual fixtures," *IEEE Transactions on Robotics*, vol. 20, no. 6, pp. 953–966, 2004.

[8] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, K. V. Alwala, N. Carion, C.-Y. Wu, R. Girshick, P. Dollár, and C. Feichtenhofer, "Sam 2: Segment anything in images and videos," *arXiv preprint arXiv:2408.00714*, 2024.

[9] D. Wu, H. Wang, Z. Hu, and F. Nie, "Improved deep metric learning with local neighborhood component analysis," *Information Sciences*, vol. 617, pp. 165–176, 2022.

[10] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," *arXiv preprint arXiv:1706.06083*, 2017.

A Tactile Sensing Array for Force Quantification in Robot Assisted Surgery

C. DeLorey[1], M.Brockdorff[1], A. Metcalf[1], B. Calmé[1], P. Valdastri[1], D. Jones[1]

[1]*School of Electrical and Electronic Engineering, University of Leeds, elchd@leeds.ac.uk, D.P.Jones@leeds.ac.uk*

INTRODUCTION

Robotic minimally invasive surgery has transformed clinical practice, however as surgeons are mechanically separated from their instruments, this removes natural force feedback. In open and laparoscopic surgery, tactile cues guide safe tissue handling [1], but RAMIS systems force surgeons to rely mainly on vision [2]. Misjudged force can cause ischemia from excessive pressure or tissue slip from insufficient grip. On-jaw sensing is especially important for monitoring contact pressures, shear, and friction in variable surgical environments. Providing force feedback has been shown to reduce excessive force use and shorten procedure times [3].

As surgical robotics trends towards sensorisation and automation [4], achieving force feedback in RAMIS systems could enable enhanced robot autonomy. Prior work has explored several transduction mechanisms for instrumented surgical graspers, however few achieve force sensing at the tool-tissue interface. This work presents a compact force sensor with three sensing elements (taxels) affixed to the grasping face of a *da Vinci* surgical tool. The sensor features a miniaturised design using Hall effect sensors and magnets embedded in a deformable elastomer contact layer. Validation is performed on a silicone tissue phantom and ex vivo porcine tissue, demonstrating force sensing during soft material manipulation.

MATERIALS AND METHODS

In this work, magnet-based force sensing is adopted for its compact sensor packages, electrical simplicity, high sampling rates, however requiring compensation to mitigate stray-field and orientation biases [5]. The sensor consists of three 3-DoF taxels, each 2.50 x 1.2 x 2.50 mm, on removable mounting affixed to the grasping face of a *da Vinci* fenestrated bipolar forceps jaw (Fig. 1i).

Manufacturing

A custom PCB with eight 3-axis Hall effect sensors (AK09973D, Asahi Kasei Microdevices, Japan) arranged in four pairs was made. Two permanent magnets (0.5 x 0.5 x 0.2 mm; SuperMagnetMan, USA) for each taxel are placed within the elastomeric layer mould, held in position by a single larger permanent magnet on the exterior of the mould. Rigid NdFeB magnets are used for their high remanence, giving high magnetic field signal at small scale. EcoFlex 00-30 elastomer (Smooth-On Inc., USA) is poured into the mould, then cured at 40°C. This elastomer skin is removed, with the magnets encased inside (Fig. 1ii-iii). This sensor skin and PCB (Fig. 1iv) are affixed to a 3D-printed mounting (Fig. 1v).

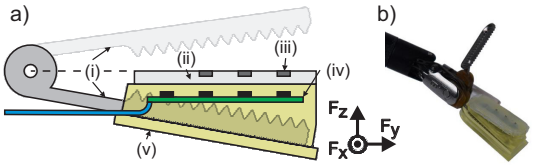


Fig. 1 (a) Schematic section view: (i) original grasper jaw, (ii) elastomer layer, (iii) taxel magnet, (iv) PCB, (v) 3D-printed mount; (b) sensor mounted in fenestrated forceps.

Calibration

The sensor data is acquired from a microcontroller over serial connection at 500 Hz, recorded via a PC. To account for stray magnetic fields as the sensor changes orientation, the average measurement from the most proximal pair of sensors is subtracted from the each taxel's respective field measurement.

Each taxel on the sensor is calibrated separately using a parallel robot platform to apply precise deformations to each taxel. A 6-axis load cell (Nano 17, ATI, USA) is used to record the applied forces and torques. The taxels were indented along multiple 3D vectors, forming a concentric cone trajectory. The calibration data is synchronised to a rate of 150 Hz, totalling ∼400,000 datapoints per taxel calibration. Load cell forces were filtered with a 3rd order Butterworth filter, (1 kHz, 10 Hz cutoff) to remove high frequency noise. The calibration raw magnetic field values and load cell forces are used to train a Feed-Forward Neural Network (1 ReLU hidden layer, 25 neurons). The calibrations yield prediction RMSE (mean [std]) of 0.954 kPa [0.436 mN], 0.756 kPa [0.033], and 4.688 kPa [0.759] in the X, Y, and Z axes respectively. Hysteresis testing shows percent signal errors: Taxel 1, 14.75% [0.68]; Taxel 2, 10.67% [0.39]; Taxel 3, 13.55% [1.17], in line with the literature.

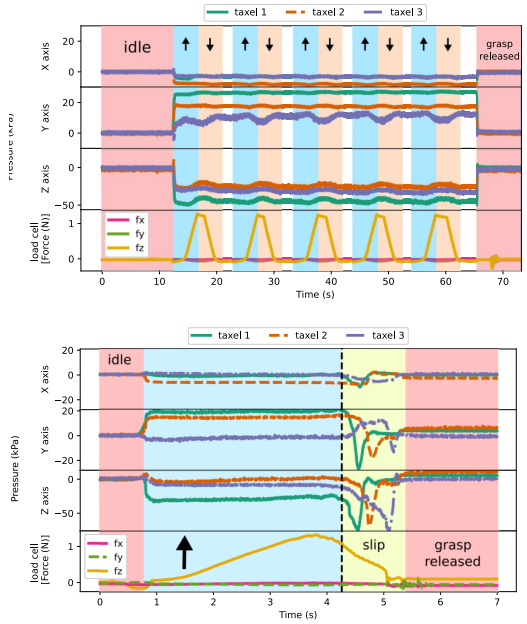


Fig. 2 Retraction (top) and slip (bottom) experiments on silicone tissue simulant. Dashed vertical line denotes slip event.

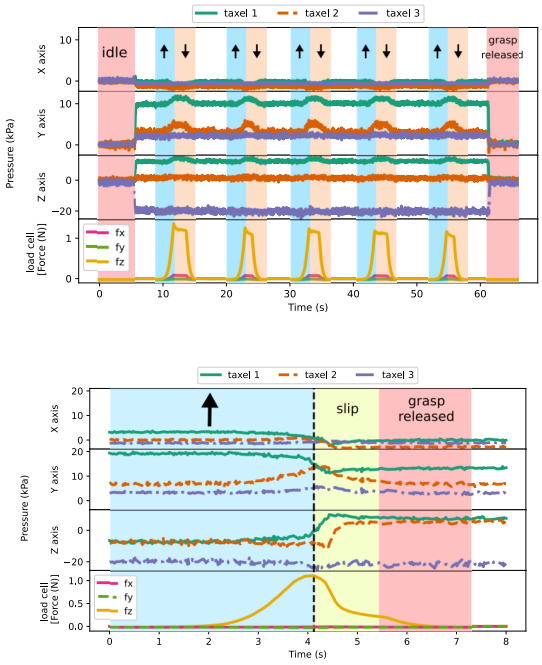


Fig. 3 Retraction (top) and slip (bottom) experiments on porcine tissue. Dashed vertical line denote slip event.

Experiments

The sensor is affixed to one jaw of a fenestrated bipolar forceps instrument for the *da Vinci* surgical robot, controlled via the da Vinci Research Kit [6]. A sample is affixed to a 6-DoF load cell (Nano17, ATI, USA) for data collection. To improve sensor readings, a 3D-printed toothed jaw part is affixed to the non-sensorised grasper jaw. Cyclic retraction tests and slip tests were performed on silicone tissue simulant and ex-vivo porcine small intestine tissue, ethically acquired from the food chain. These porcine samples were stored in phosphate-buffered saline to preserve hydration and mechanical properties.

RESULTS

Silicone simulant retractions show cyclic applied forces in X, Y, Z as the material is pulled (Fig. 2). Minor pressures in X (~-8 kPa) are likely due to alignment variations of the sensors and the load cell.

Slip experiments on silicone showed force change and slip at sequential taxels, with the slip propagating distally along the sensor (vertical black dashed line, Fig. 3).

Tissue retraction yielded peaks in pull force primarily in Y, then Z axes. The sensors are able to detect this change (Fig. 3). Retraction is performed at 10 mm and then at 15 mm to observe effect of greater pull force. The sensor registered a ~2.5x increase in max pull force (mean [std] : 29.592 kPa [3.974 kPa] versus 74.49 kPa [11.939 kPa]), maintaining grasp and simultaneously measuring shear force. In tissue slip experiments (Fig. 3) we see a similar hyperelastic response, as the tissue slips in the jaws.

DISCUSSION

The Hall-effect force sensor presented can measure pressures in the 0–255 kPa range, typical of RAMIS tasks. Tissue properties (thickness, hydration, elasticity, etc.) and out-of-plane grasping leads to variable grasp behaviours. The retraction and slip experiments focused on the Y axis, the principal axis of slip in these experiments, a simplified demonstration. Tissue retraction showed ~50% lower pressures than silicone, and slip experiments had less pronounced slip events, pointing to inconsistent contact, heterogeneous properties, or stick–slip behaviour. This sensor enables slim 3-DoF force sensing on a surgical robot instrument. Future work will explore multi-axis slip detection, and integration with learning agents for automation of tissue manipulation.

REFERENCES

[1] G. Tholey, J. P. Desai, and A. E. Castellanos, "Force feedback plays a significant role in minimally invasive surgery: results and analysis," *Annals of surgery*, vol. 241, no. 1, p. 102, 2005.

[2] R. V. Patel, S. F. Atashzar, and M. Tavakoli, "Haptic Feedback and Force-Based Teleoperation in Surgical Robotics," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 1012–1027, 7 2022.

[3] A. Abiri, J. Pensa, A. Tao, J. Ma, Y.-Y. Juo, S. J. Askari, J. Bisley, J. Rosen, E. P. Dutson, and W. S. Grundfest, "Multi-Modal Haptic Feedback for Grip Force Reduction in Robotic Surgery," *Scientific Reports*, vol. 9, no. 1, p. 5016, 3 2019.

[4] S. Schmidgall, J. D. Opfermann, J. W. Kim, and A. Krieger, "Will your next surgeon be a robot? Autonomy and AI in robotic surgery," Tech. Rep., 2025. [Online]. Available: https://www.science.org

[5] I. Waters, D. Jones, A. Alazmani, and P. Culmer, "Utilising incipient slip for grasping automation in robot assisted surgery," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1071–1078, 2021.

[6] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da Vinci® Surgical System," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5 2014, pp. 6434–6439.

An electromagnetic catheter for on-demand release and in-situ manipulation of magnetic therapeutic cargoes

Riccardo Panizzolo[1], Tamerlan Srymbetov[1], Arianna Menciassi[1], and Veronica Iacovacci[1]

[1]*The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy*

INTRODUCTION

Magnetic manipulation and wireless actuation are gaining growing interest in biomedical applications, as magnetic fields are harmless to the human body and enable minimally invasive procedures. Among these applications, targeted drug delivery represents an innovative approach that aims to minimize systemic drug exposure while enhancing the therapeutic effect by directing the drug precisely to the pathological site. In this context, the use of magnetic microagents (MMAs) has grown significantly due to their versatility and controllability. Magnetic agents, such as magnetic microparticles and microbeads, are a particularly versatile embodiment of this concept [1]. Magnetic actuation can be performed with permanent magnets and/or electromagnets: the latter, being current-driven, allow finer and on-demand modulation of fields and gradients, and have enabled the wireless steering of dexterous MMAs in hard-to-reach anatomical sites. Most existing platforms rely on remote, large-workspace electromagnetic systems that steer a single agent at a time, while local actuation strategies closer to the target and capable of multi-agent or multi-stage control remain comparatively unexplored. This becomes limiting in drug delivery: while several solutions have been proposed to transport and eject the therapeutic cargo, active control of the released particles is still largely missing, leaving them prone to dispersion and sedimentation before reaching the target. A first step in this direction is represented by MILiMAC, a flexible catheter for single microbead steering [2]. Extending this concept, opportunities remain in broadening local actuation toward integrated cargo release, multi-agent control, and operation in unconstrained workspaces. In this work, we present an electromagnetic catheter designed to integrate three functionalities within a single device: (i) catheter navigation toward the lesion site, (ii) on-demand cargo release, and (iii) active manipulation of the released cargo. Functionalities (i) and (ii) build on a design previously reported by the authors [3], while the main contribution of the present work is functionality (iii): an electromagnetic actuation unit and the associated control strategies enabling post-release cargo manipulation, so as to counteract dispersion and promote local drug mixing. Accordingly, the results reported here characterize post-release steering, aggregation, and rotation in place in a benchtop workspace.

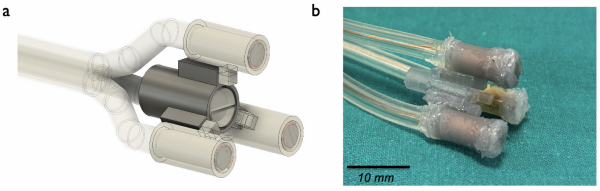


Fig. 1: Design and implementation of the proposed catheter: a) Concept of the whole system integrating the release unit and the control unit ; b) 3D-printed final prototype with silicone tubes for coils' wires and seals for fluidic media.

MATERIALS AND METHODS

The catheter includes two main blocks: a navigation/release unit and a local control unit. The first leverages remote magnetic control for navigation and on-demand cargo release [3]. The latter, here presented, is intended for active manipulation of the released cargo and includes three miniaturized coils, radially arranged around the carrier at 120° from each other (*Fig.* 1) . This configuration represents the minimum number of coils required to produce a rotating magnetic field in the working plane through phase-shifted sinusoidal currents. The coils are made of four layers of copper wire wrapped around a ferrite core (15 mm long, 1.5 mm in diameter). Empirical assessment allowed us to establish that at least 15 mT at the coil tip are needed for bead manipulation. FEM assisted analysis allowed to dimension the coils to support this magnetic output. The structure holding the coils and the carrier was 3D printed using a BioMed Clear resin. Furthermore, silicone tubes connected directly to the coil holders were used to seal the wires from the medium. Different MMAs were tested as cargo. Microbeads were fabricated by a precipitation method, embedding iron-oxide or neodymium-iron-boron microparticles in a highly viscous solution coated with a PTFE shell. The resulting beads had outer diameters in the range of 200-800 μm. A LabVIEW-based real-time control system was developed to provide a user interface for precise coil current control, facilitating the implementation of pre-planned trajectories. The operator can choose between two actuation modes: phase-shifted sinusoidal currents producing rotating fields for circular bead steering, and sequential activation of single coils with constant currents, exploiting the local gradient at each coil tip to drive the beads along triangular paths. The connection between the software and the hardware was established using a data acquisition card *(DAQ; National*

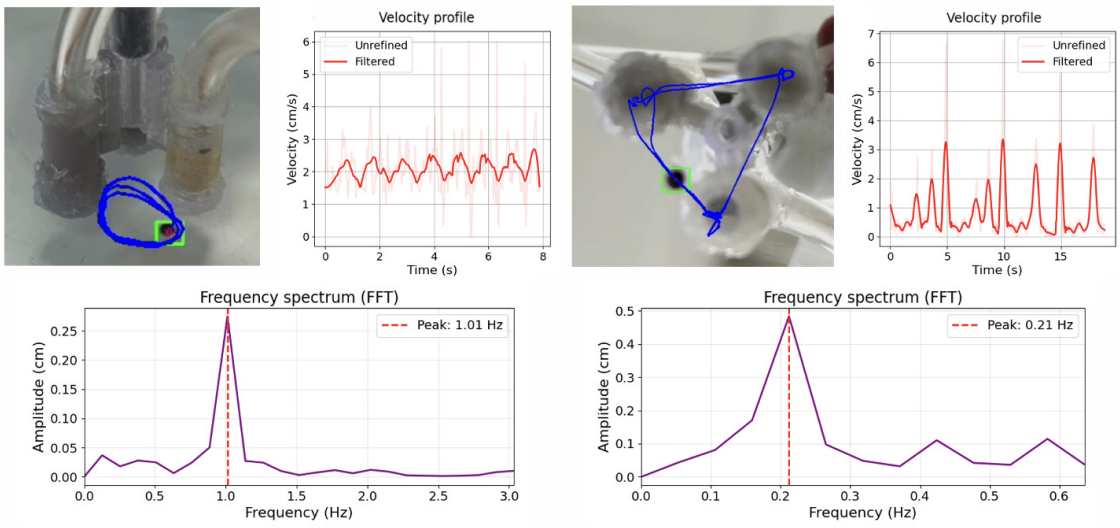


Fig. 2: Trajectory of the microbead in water within the workspace, filtered velocity, and frequency spectrum of the bead motion, using the two different control strategies. Left column: Sinusoidal currents (circular trajectory). Right column: sequential coil activation (triangular trajectory). Top row: extracted trajectory and velocity profile; bottom row: corresponding frequency spectrum via FFT.

Instruments, USB 6009). To command the current through each coil, three current drivers were employed *(NPS-090-10-Z, Copley Controls)*, connected in turn to an *AC-DC* power supply. The drivers were configured with a unitary V to A gain, keeping the operating currents below 2 A and well within the DAQ output range. Furthermore, to parametrize and analyze the resulting microbead trajectories, a computer vision algorithm was used to reconstruct offline trajectory and velocity data from video recordings, correlating them with the frequency. The beads were tracked frame-by-frame with a CSRT correlation-filter tracker, the instantaneous velocity was then obtained by finite differences and smoothed with a Savitzky-Golay filter. A frequency-domain analysis (FFT) of the velocity signal was then performed to extract the dominant motion frequency, which was correlated with the input frequency.

RESULTS

The catheter defines a cylindrical workspace of approximately 10 mm in diameter in which the post-release control of the cargo takes place. The system was validated in both water and silicone oil, demonstrating its ability to release the cargo from the carrier and subsequently steer the released microbeads through the local control unit. Sinusoidal currents applied to the coils generated rotating magnetic fields that drove individual microbeads along circular trajectories, while sequential activation of single coils with constant currents and low frequencies produced triangular paths between the three coil tips. The computer vision algorithm confirmed that the dominant motion frequency of the bead matches the input current frequency, and that the orbit radius decreases with increasing rotation frequency, consistent with the rotating-field actuation principle (*Fig.* 2). When several microbeads were dispersed in the workspace, the same actuation strategies promoted dipole-dipole interactions that led to the formation of chains and larger aggregates, which could in turn be steered as a single entity along the same preplanned trajectories. Above approximately 5 Hz for both individual beads and aggregates, in-place rotation was observed. In this regime the agents rotate around their own axis without net translation, providing an additional mechanism to induce localized fluid agitation and promote drug mixing at the target site. The actuation current was clipped at 2 A to prevent coil overheating, and the drive frequency was hardware-limited to 20 Hz.

CONCLUSIONS AND DISCUSSION

We presented an electromagnet-based catheter that combines drug release with active post-release control of the cargo, enhancing drug-mixing capabilities and addressing the dispersion and sedimentation issues that limit conventional release mechanisms. The proposed control strategy was successfully validated for single-bead steering and multi-bead aggregation, immersed in different media. Future work will address real-time localization compatible with clinical imaging and quantitative characterization of the mixing efficiency, exploiting the rotating magnetic field also for vortex-like swarm formation.

REFERENCES

[1] V. Iacovacci, E. Diller, D. Ahmed, and A. Menciassi, "Medical microrobots," *Annual Review of Biomedical Engineering*, vol. 26, pp. 561-591, 2024.

[2] J. Sikorski, S. Mohanty, and S. Misra, "Milimac: Flexible catheter with miniaturized electromagnets as a small-footprint system for microrobotic tasks," *IEEE Robotics and Automation Letters*, vol. 5, pp. 5260-5267, 2020.

[3] T. Srymbetov, G. D. Angelis, A. Menciassi, and V. Iacovacci, "Millimeter-scale magnetic carrier for on-demand delivery of magnetic and non-magnetic microparticles suspensions," *International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 2024.

Selectively Triggerable Biopsy Capsule via Multi-stage Magnetic Torsion Springs

Simon Frieler[1], Jean-Paul P.M. de Vries[3], Sarthak Misra[1], Eric Diller[2], and Venkatasubramanian Kalpathy Venkiteswaran [1]

[1]Department of Biomechanical Engineering, University of Twente, The Netherlands
[2]Department of Mechanical and Industrial Engineering, University of Toronto, Canada
[3]Department of Surgery, University Medical Center Groningen, The Netherlands

INTRODUCTION

Endoscopy remains the gold standard for gastrointestinal (GI) tract diagnostics; however, it is associated with patient discomfort and psychological stress [1]. In response, ingestible capsule robots have emerged as a promising alternative due to their miniature size.

Magnetic actuation presents a popular navigation strategy for capsule robots, offering untethered operation, compatibility with biological environments, and precise control, enabling multi-axis manipulation. One of the most common functions needed for capsules to be used in GI tumor detection is the ability to perform *in situ* tissue biopsy. Multi-barrel capsules compactly accommodate multiple tools. Existing six-tool designs are either motorized and tethered for power [2] or wireless but couple selection and deployment, imposing a fixed sampling sequence [3]. Thus, wireless independent control of tool selection and deployment remains undemonstrated.

This study introduces a multistage magnetically actuated biopsy capsule with a cam mechanism that decouples tool selection from deployment, enabling programmable sampling sequences. Magnetic actuation enables both capsule navigation and wireless biopsy actuation. The capsule can generate deployment forces of up to 1.5 N, sufficient to puncture tissue in the GI tract. *Ex vivo* trials on a porcine stomach show rolling and reorientation as locomotion strategies, enabling the successful collection of 1 mg of biological tissue and 15 µL of simulated gastric juice.

MATERIALS AND METHODS

In a clinical workflow, the patient would swallow the capsule with water, allowing peristalsis to carry it through the esophagus to the stomach (Fig. 1 (a)). An electromagnetic actuation system then applies an external magnetic field. Upon entering the GI tract, the capsule can be manipulated and triggered by applying magnetic fields of different orientations and magnitudes (Fig. 1 (b,c)). A rotating magnetic field of 5 mT induces a rolling motion that enables precise navigation across uneven terrain such as gastric rugae (Fig. 1 (b)). Applying a magnetic field of 5 mT perpendicular to the capsule's longitudinal axis enables reorientation of the capsule (Fig. 1 (c)). Once the target site has been reached, multiple liquid and tissue biopsies can be triggered magnetically (Fig. 1(d,e)).

The capsule comprises a selection module and a deployment module, both of which contain a magnetic torsion spring (MTS) mounted on a central guide shaft. MTS_r

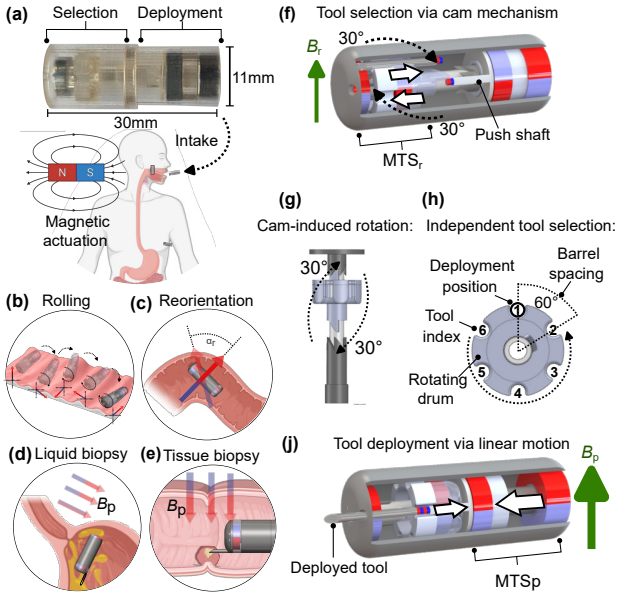


Fig. 1: (a) Physical prototype of ingestible biopsy capsule (b-c) Magnetically-induced rolling and reorientation as locomotion strategy (d-e) Magnetically-triggered liquid and tissue biopsy (f) Tool selection via cam mechanism and magnetic torsion spring (g-h) Drum rotation mechanism (j) Tool deployment via linear motion generated by second magnetic torsion spring.

drives the rotation of the drum containing the biopsy tools (Fig. 1 (f)). MTS_r triggers at a magnetic field strength of B_r = 18 mT, causing a magnet inside the drum to align with B_r (Fig. 1 (f)). Consequently, the magnet rotates relative to the fixed counter-magnet, switching their interaction from attraction to repulsion. This causes the drum to translate toward MTS_r, while an integrated cam mechanism induces a 30° rotation (Fig. 1 (g)). This motion aligns the drum containing the biopsy tool with the push shaft of MTS_p (Fig. 1 (h)). When the field increases from B_r to B_p = 38 mT, MTS_p triggers, and the push shaft deploys the biopsy tool out of the capsule with a maximum stroke length of 5 mm (Fig. 1 (j)). During this process, MTS_r remains in its triggered state. When the field is gradually removed, the sequence reverses. First, MTS_p collapses, retracting the biopsy tool from the target and storing it within the drum. Next, MTS_r collapses, returning the drum to its initial position.

Selective indexing of the biopsy tool is achieved by applying a magnetic field B_r, which induces a 30° rotation of the drum. Upon removal of the field, the drum attracts

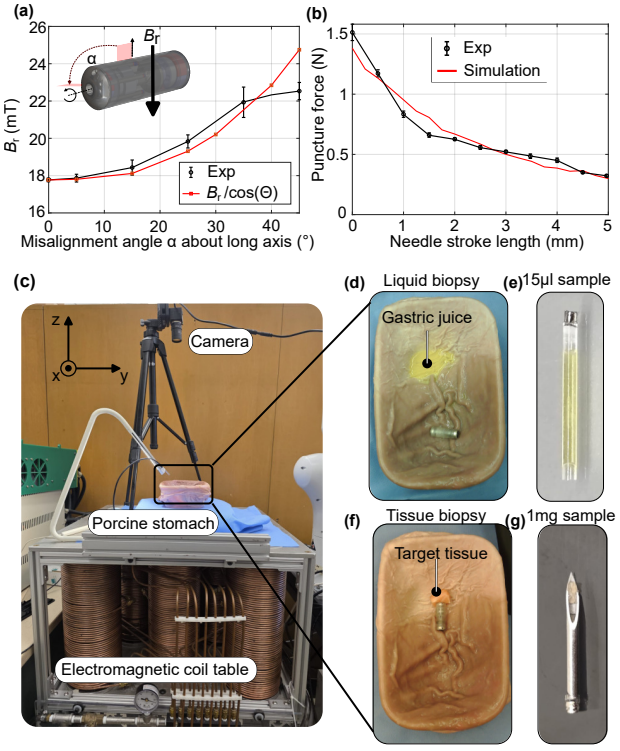


Fig. 2: (a) Required field strength B_r (mean±SD) versus misalignment angle (α) (b) Puncture force (mean±SD) versus needle stroke (c) Experimental setup: porcine stomach on electromagnetic coil table (d-e) Liquid biopsy of simulated gastric juice (f-g) Tissue biopsy of biological tissue.

to the fixed magnet, and the cam mechanism induces a further 30° rotation. Consequently, each actuation cycle (application and release of B_r) induces a total rotation of 60°, corresponding to the angular spacing between adjacent barrels of the drum (Fig. 1 (h)). This enables an independent tool selection with each actuation cycle.

To ensure operation within the limits of the actuation system, the trigger thresholds of MTSr and MTSp are designed to be 20 mT and 40 mT, respectively, using magnetostatic finite element analysis (FEA) in COMSOL Multiphysics 6.0. The deployment force is characterized by mounting the capsule on a linear stage and positioning it at varying distances (0-5 mm) from a load cell (GSO-500, Transducer Techniques, USA). The capsule is actuated three times at each distance, and the peak force is recorded. In addition, the deployment force of MTS_p is analyzed using magnetostatic FEA. To investigate jamming and friction caused by shear forces resulting from a misalignment about the long axis with respect to B_r, the capsule is rotated by $\alpha \in [0, 45]^\circ$ in a static field, and the trigger threshold is measured after three trials.

In experiments, the capsule is actuated using an unobtrusive magnetic navigation system, capable of generating non-uniform magnetic fields up to 38 mT in the x- and y-axes and 47 mT in the z-axis at a working distance of 120 mm [4] (Fig. 2 (c)). A 19 G infusion needle and a capillary suction tube (1 mm in diameter) are used for tissue and liquid biopsy. The experimental setup consists of a porcine stomach conformed to a container and suspended 120 mm above the magnetic actuation system, with biological tissue and simulated gastric juice (pH-value 1.2) serving as targets.

RESULTS

Capsule specifications are listed in Table I. Experiments yield trigger thresholds of $B_r = 18 \pm 0.19$ mT and $B_p = 38 \pm 0.30$ mT (Table I). For misalignment angles $\alpha \in [0, 45]^\circ$, the trigger threshold for MTS_r increases from 18 ± 0.19 mT to 22 ± 0.92 mT (Fig. 2 (a)). The needle deployment force decreases with increasing stroke length, reaching a maximum of 1.5 ± 0.14 N at 0 mm stroke and decreasing to 0.25 ± 0.04 N at 5 mm (Fig. 2 (b)).

Ex vivo trials on the porcine stomach show that the capsule leverages rolling and reorientation as locomotion strategies to reach the target (Fig. 2 (d,f)). This enables the successful sampling of 1 mg of biological tissue and 15 µL of simulated gastric juice using a hollow needle and a capillary suction tube, respectively (Fig. 2 (e,g).

TABLE I: Overview of the capsule specifications.

Specification	**Value**
Dimensions (mm)	11x30
Weight (g)	35
Maximum deployment force / *simulated* (N)	1.5 ±0.14 / *1.38*
Trigger threshold MTS_r experimental / *simulated* (mT)	18 ±0.19 / *20*
Trigger threshold MTS_p experimental / *simulated* (mT)	38 ±0.30 / *40*
Collected tissue sample (mg)	1
Collected liquid sample (µL)	15

DISCUSSION AND CONCLUSION

At 11×30 mm (2.85 cm^3), the proposed capsule remains below the reported swallowability limit of 3 cm^3 [5], and is comparable in size to existing multi-tool capsule robots [2], [3], while enabling independent tool selection and deployment . The inner diameter of 2.5 mm allows integration of a camera module for real-time visual feedback during use. The maximum deployment force of 1.5 N with a 19 G needle is high enough to penetrate GI tissue. Beyond liquid and tissue biopsies, the capsule's use could be extended to targeted drug delivery. However, challenges related to cross-contamination and capsule visualization in an insufflated GI tract during gastroscopy remain critical for clinical use.

REFERENCES

[1] S. N. Adler and Y. C. Metzger, "Pillcam colon capsule endoscopy: recent advances and new insights," *Ther. Adv. Gastroenterol.*, vol. 4, no. 4, pp. 265–268, 2011.

[2] Z. Li, T. Wei, Y. Zhang, Q. Liu, C. Du, C. Shi, and C. Hu, "Rotating multibarreled capsule robot for multiple biopsies and on-demand drug delivery," *IEEE/ASME Transactions on Mechatronics*, vol. 30, no. 6, pp. 7089–7099, 2025.

[3] S. Park, M. C. Hoang, J. Kim, and S. Park, "Multiple sampling capsule robot for studying gut microbiome," *Advanced Intelligent Systems*, vol. 7, no. 11, p. 2300625, 2025.

[4] A. Schonewille, C. He, C. Forbrigger, N. Wu, J. Drake, T. Looi, and E. Diller, "Electromagnets under the table: An unobtrusive magnetic navigation system for microsurgery," *IEEE Trans. Med. Robot. Bionics*, vol. 6, no. 3, pp. 980–991, 2024.

[5] V. H. Le, M. C. Hoang, V. D. Nguyen, E. Choi, C.-S. Kim, J.-O. Park, J. Kim, and B. Kang, "A novel biopsy capsule endoscope for wireless intestinal tissue collection," in *2018 International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS)*, Nagoya, Japan, 2018, pp. 1–6.

Volumetric Magnetic Model and Polar-Coordinate Inverse Solver for Model-Based Control of a Magnetically Actuated Capsule Robot

Vito Romanello[1], Tamerlan Srymbetov[1], Arianna Menciassi[1], and Veronica Iacovacci[1]

[1]*The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy*
vito.romanello@santannapisa.it

INTRODUCTION

Magnetically actuated capsule robots are a promising platform for minimally invasive interventions such as targeted drug delivery and soft-tissue inspection [1]. Closed-loop position control of these untethered devices relies on an accurate model of the interaction between an external permanent magnet (EPM) and the internal permanent magnets (IPMs) in the capsule.

Most existing controllers rely on the classical point-dipole approximation, which loses accuracy when the working distance is comparable to or lower than the magnet size—a regime unavoidable in compact biomedical setups [2], [3]. This error is inherited by any dipole-based controller. A separate limitation affects model-based iterative controllers (MBICs): when formulated in Cartesian coordinates, the axisymmetric EPM geometry creates near-degenerate solutions along the magnetisation axis, increasing the iterations required for convergence.

This work addresses both limitations. First, the point-dipole approximation is replaced by an exact volumetric model of finite-size cylindrical magnets, valid throughout the working volume, singularity-free, and characterized here in simulation. Second, the MBIC inverse problem is reformulated in polar coordinates (r, θ, z) centred on the capsule, which align a search direction with the azimuthal soft mode. The same direction-aware warm start $\theta_0 = \mathrm{atan2}(d_y, d_x)$ initialises both formulations at the same EPM pose, so the reduction in iteration count isolates the effect of the parametrisation.

MATERIALS AND METHODS

A. Experimental platform

The platform consists of a Doosan M0609 6-DOF arm carrying a NdFeB cylindrical EPM with a coaxial bore (outer/bore ⌀60/⌀10×70 mm, axial magnetisation, grade N42, B_r=1.28 T). The EPM axis is fixed along the world z-direction with yaw in the xy-plane and zero pitch, consistent with the state of the art [2]. The working height (z=78 mm) is set by the pull-out limit, keeping the vertical magnetic force below the capsule weight to prevent lift-off. The controlled object is a 77 mg dual-IPM capsule (9×4 mm, B_r=1.08 T, two axially counter-opposed IPMs), immersed in a water-filled tank [4]. The setup and capsule free-body diagram are shown in Fig. 1. All results reported here are obtained in software-in-the-loop (SIL): the control code runs against a simulated robot–capsule pair rather than the physical setup.

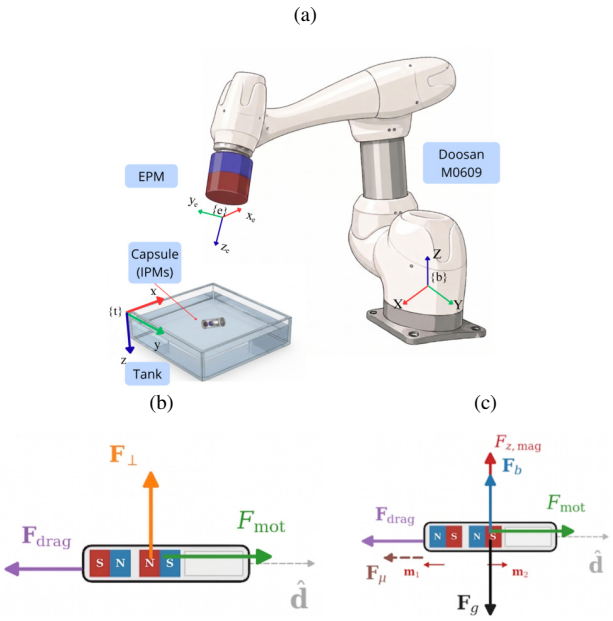


Fig. 1: Setup and capsule free-body diagram. (a) Base $\{b\}$, EPM $\{e\}$ and tank $\{t\}$ frames. (b) In-plane forces (top view, xy): propulsion F_{mot} along $\hat{\mathbf{d}}$, lateral residual $\mathbf{F}_\perp$ and viscous drag $\mathbf{F}_{\mathrm{drag}}$. (c) Vertical balance (side view, xz): gravity $\mathbf{F}_g$, buoyancy $\mathbf{F}_b$ and the magnetic $F_{z,\mathrm{mag}}$, kept below the pull-out limit to prevent lift-off.

B. Magnetic model

The force acting on the capsule, required both to predict the propulsion delivered by a candidate EPM pose and to define the cost in eq. (3), is computed as:

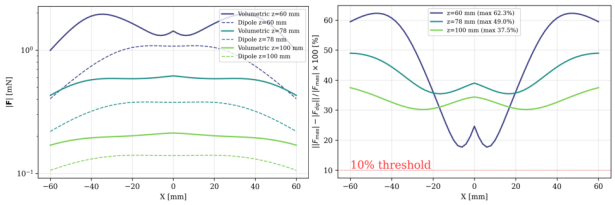
$$\mathbf{F}_i = \mu_0 \left(\mathbf{J}^{\mathrm{ext}} - \mathbf{J}^{\mathrm{int}}\right) \mathbf{m}_i, \tag{1}$$

where $\mathbf{J}^{\mathrm{ext}}$ and $\mathbf{J}^{\mathrm{int}}$ [5] are the exact 3×3 field-gradient matrices of the EPM outer cylinder and of its central bore. Their difference models the bored cylinder by superposition, evaluated at the IPM centre, and $\mathbf{m}_i$ is the IPM moment. Unlike the dipole approximation, eq. (1) is valid at arbitrary working distances and is free of near-field singularities (Fig. 2).

C. Controllers

Three controllers are compared on the same trajectory: a model-free *spatial PD* as baseline, the *MBIC* with the proposed polar reformulation, and a *Cartesian MBIC* variant as control condition.

The spatial PD maps tracking error into a robot offset,



(a) Force magnitude (log scale) (b) Relative error

Fig. 2: Volumetric vs. point-dipole [5]: lateral scan in x over 120 mm at three fixed heights z (corresponding to the x-axis of the world frame in Fig. 1a). (a) Force magnitude (log scale): volumetric (solid) and point-dipole (dashed) at z=60, 78, 100 mm; (b) relative error; dashed line marks the conventional 10 % threshold.

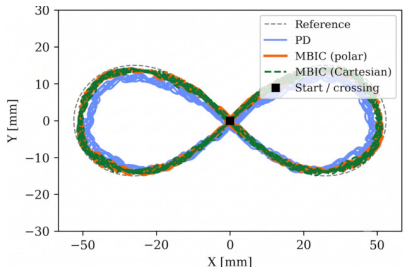


Fig. 3: Closed-loop tracking of the lemniscate (SIL): capsule trajectories in the xy-plane.

clamped to a maximum displacement $d_{\max}$:

$$\mathbf{p}_{\text{rob}} = \mathbf{p}_{\text{cap}} + \text{sat}(k_p\mathbf{e} + k_d\dot{\mathbf{e}},\ d_{\max}), \tag{2}$$

with k_p=0.5, k_d=0.05 s and $d_{\max}$=40 mm (empirical). The MBIC [2] uses an outer loop (position error → desired velocity, $v_{\max}$=20 mm/s, dead-zone 1.5 mm) and an inner loop (drag model → required force) that selects the EPM pose by solving:

$$\min_{r,\,\theta,\,z} \left|F_{\text{mot}} - F_{\text{req}}\right| + w_\perp \|\mathbf{F}_\perp\| + w_z p_z + w_d p_d, \tag{3}$$

where $F_{\text{mot}}=\mathbf{F}_i\cdot\hat{\mathbf{d}}$ with $\mathbf{F}_i$ as in eq. (1), $\mathbf{F}_\perp$ is the lateral residual, p_z, p_d penalise lift-off and unfavourable EPM directions, and $w_\perp, w_z, w_d$>0 are fixed weights. The Cartesian variant uses the same cost and model but replaces the polar parametrisation with a Cartesian (x, y, z) search using the same directional warm start $\theta_0 = \text{atan2}(d_y, d_x)$, so the identical warm start isolates the coordinate parametrisation.

D. Simulation framework

All results are from the SIL framework in Python (f_{sim}=100 Hz, f_{ctrl}=10 Hz, T=40 s) comparing the three controllers under identical conditions. We report near-path RMSE (cross-track distance to the reference) and coverage (waypoints within 4 mm), mean and RMS of $\|\mathbf{F}_\perp\|$, and the average Sequential Least-Squares Quadratic Programming (SLSQP) iterations per control step.

E. Reference trajectory

A Lemniscate of Bernoulli (a=30 mm, 300 waypoints) is the benchmark: its varying curvature and self-intersection reversal span the full heading range $\theta\in[0, 2\pi)$, stressing the polar warm start.

RESULTS

Magnetic model. The dipole's relative error spans 36–49 % at the nominal distance z=78 mm (1.3 EPM diameters) and peaks at 62 % at z=60 mm (Fig. 2); being pose-dependent, it cannot be absorbed by a scalar calibration.

TABLE I: Controller comparison on the lemniscate (SIL).

Metric	PD	MBIC (Polar)	MBIC (Cartesian)
Near-path RMSE [mm]	2.62	1.21	1.22
$F_\perp$ mean [mN]	0.157	0.129	0.129
$F_\perp$ RMS [mN]	0.194	0.181	0.181
SLSQP iter. (avg)	—	4.0	15.0

Closed-loop tracking. Fig. 3 shows the three tracked trajectories and Table I the corresponding metrics. With identical model and warm start in both MBIC variants, the polar parametrisation cuts the average SLSQP iterations from 15 to 4 at identical accuracy and lateral force, diagonalising the azimuthal soft mode of the axisymmetric field. The lower iteration count leaves real-time margin for the higher rates and 3-D pose search of the hardware stage, where the Cartesian degeneracy would otherwise dominate the cost. Compared to the spatial PD, MBIC reduces the near-path error by ≈ 54 % (1.21 vs 2.62 mm) and the mean lateral force by ≈ 18 % (0.129 vs 0.157 mN). All three controllers reach 100 % coverage.

CONCLUSIONS AND DISCUSSION

The volumetric model is worth its complexity when its error is the dominant one. Two error sources compete.

The dipole error is systematic: for a given geometry and pose it is fixed and repeatable (up to ≈49 % here). It acts as a fixed bias inside the inverse map, and the volumetric model removes it exactly. Environmental disturbances—ferrous objects, hydrodynamics—are instead additive and time-varying, but at the capsule they stay a small fraction of the EPM field, because the EPM is the closest and strongest source.

What this is worth depends on the controller. In feed-forward the model sets the commanded force, so a 49 % model error becomes a 49 % force error. Under feedback the outer loop corrects slow errors, and using a model pays off as reduced control effort: the model-based MBIC needs 18 % less lateral force and 54 % less near-path error than the model-free PD (Table I). An accurate model matters most in compact setups, where the magnet—not the environment—sets the field.

Future work will add hardware-in-the-loop validation and deployment with ultrasound localisation.

REFERENCES

[1] Z. Wu, X. Wang, Y. Lu, and Q. Xu, "Multichamber magnetic capsule robot for selective liquid sampling and drug delivery," *National Science Review*, vol. 12, no. 11, p. nwaf400, 2025.
[2] G. Pittiglio, L. Barducci, J. W. Martin, J. C. Norton, C. A. Avizzano, K. L. Obstein, and P. Valdastri, "Magnetic levitation for soft-tethered capsule colonoscopy actuated with a single permanent magnet: A dynamic control approach," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1224–1231, 2019.
[3] A. W. Mahoney and J. J. Abbott, "Generating rotating magnetic fields with a single permanent magnet for propulsion of untethered magnetic devices in a lumen," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 411–420, 2014.
[4] T. Srymbetov, G. De Angelis, L. Cinus, A. Menciassi, and V. Iacovacci, "Push type magnetic milli-scale robotic carriers for controlled cargo delivery," *IEEE Transactions on Mechatronics*, 2026, under review.
[5] F. Masiero and E. Sinibaldi, "Exact and computationally robust solutions for cylindrical magnets systems with programmable magnetization," *Advanced Science*, vol. 10, no. 27, p. 2301033, 2023.

Endoscopic Deployment of Magnetic Soft Continuum Robots for Targeted Gastrointestinal Therapy

A. Bacchetti[1], A. Brown[1], B. Calmé[1], P. Valdastri[1], J.H. Chandler[1]

[1]*School of Electronic and Electrical Engineering, University of Leeds*

mnaba@leeds.ac.uk

INTRODUCTION

Crohn's Disease (CD) is a chronic gastrointestinal (GI) condition which is characterized by the presence of patched, deep-seated regions of inflamed or scarred tissue, known as 'skip' lesions. High recurrence rates, alongside the standalone nature of current therapeutic treatments may expose the patient to greater infection risks and recovery times [1,2]. Given these challenges, novel therapeutic treatments have considered direct injection over the mucosal lining to increase therapeutic drug uptake at the target site, whilst minimizing systemic effects commonly observed in traditional infusion-based approaches [2].

Rigid GI endoscope interventions are often utilized for onboard therapeutic delivery. However, their scale, stiffness and limited dexterity can restrict access to deep-seated skip lesions. To improve upon these challenges, magnetic soft continuum robotic designs have been proposed which are navigated via magnetic forces and torques, produced from interactions between externally generated actuating magnetic fields and magnetically active regions embedded within their structures [4]. Numerous recent designs have relied on complex geometric and embedded material heterogeneity for sufficient deflection and stiffness [3,4], however this often compromised the design's simplicity, scale and ability to house additional functional components. Creating MSCRs which readily integrate with existing endoscopic equipment can facilitate minimally invasive deployment and GI access, with the endoscope's onboard visual sensing serving as a key tool for diagnostic sensing, intra-operative positioning and monitoring of treatment delivery.

In this work, we consider an existing conical tapered MSCR [5] which allows for stable, high-angle deflections under low-strength fields, and integration of fluidic microchannels (μ-channels) for targeted therapeutic delivery (**Figure 1**). We characterize achievable deflection and stability under quasistatic rotating magnetic fields, at various deployment lengths from a rigid mock endoscope tip. We explore the robot's capabilities for targeted, magnetically guided therapeutic intervention through localized delivery over a simulated CD skip lesion within the sigmoidal colon of a live porcine model.

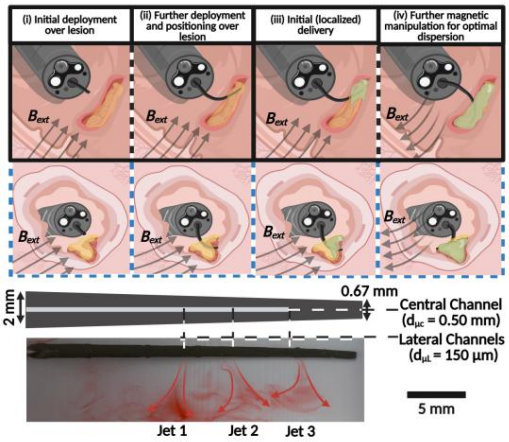


Figure 1. Conical MSCR with radial μ-channels for targeted GI delivery, and featuring an η = 3 taper ratio (proximal base diameter = 2mm, tip diameter = 0.67mm). The design facilitates (i) initial deployment, (ii) magnetically actuated positioning over the lesion, (iii) initial localized delivery, and (iv) stable high-angle deflection for optimal drug dispersion (green).

MATERIALS AND METHODS

MSCR Fabrication

An existing conical tapered MSCR design was fabricated using low-pressure injection molding. Silicone (Smooth-On Inc., Ecoflex 00-30, Pennsylvania, USA). Magnetic neodymium-iron-boron microparticles (MQFPB+ ,Magnequench GmbH, Germany) were mixed in a 1:1 mass ratio and degassed in a high-vacuum mixer at 1400 rpm, 20 kPa for 90s. Samples were cured at 40°C for 20 minutes and subsequently exposed to an axial impulse field of >4T (IM-10- 30; ASC Scientific, Rhode Island, USA). A central nitinol rod (0.50 mm diameter) was inserted axially within the case to create a central delivery channel, and connected to three radial nitinol rods (0.15mm diameter) along the MSCR's length. A 2mm nitinol rod and 0.25mm silicone fluid delivery tube were connected to the MSCR's proximal base for manual deployment and fluidic delivery.

Deflection and Stability Characterization

Deflection and stability were evaluated across 10-50 mm MSCR deployment lengths from a mock endoscope tip. Each length was exposed to a radial 5 mT magnetic field (MFG-100 magnetic field generator, Magnebotix, Zurich, Switzerland), rotated clockwise in 10° increments between 0°-360° from a vertical (0°) reference. A white tip-based marker was isolated at each field angle to identify and track tip deflections. Deflections were measured on ImageJ analysis software (National Institutes of Health, USA).

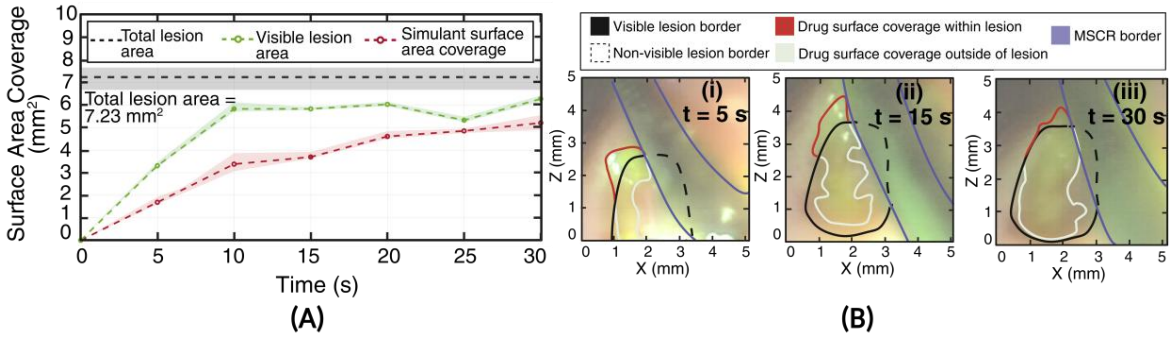


Figure 3. Surface coverage tracking and comparative delivery profiles for localized drug simulant delivery, (A) surface coverage tracking within the visible lesion area (green) and outside of the lesion border (red), (B) visual delivery profiles at: (i) 5 s, (ii) 15 s, and (iii) 15s, with annotated visible (solid black) and non-visible (dashed black) lesion borders, simulant surface coverage within the lesion's borders (light green) and external to the lesion (red), and the MSCR's border outline.

Localized Fluidic Delivery in a Live Porcine Model

The endoscope (Olympus GIF-Q260, Olympus Corporation, Tokyo, Japan) was navigated to the sigmoidal colon region via the rectum. A GI fold surface irregularity (7.23mm^2 surface area) was identified with the endoscope's onboard camera, and selected to simulate an early skip lesion. After MSCR deployment, positional ($\boldsymbol{B}$ = 2.36mT, magnetic norm $\hat{\boldsymbol{B}}_{pos}$ = [1.39 1.25 1.45]) and delivery-based magnetic fields ($\boldsymbol{B}$ = 10.34mT mT, magnetic norm $\hat{\boldsymbol{B}}_{del}$ = 4.63 6.06 6.98]). fields were applied via a single KUKA robot arm (iiwa14 LBR, KUKA, USA) with an external permanent magnet (D = 101.6mm, Grade N52). The drug simulant was fabricated by mixing a 99.995:0.005 mass ratio of distilled water to fluorescent green dye powder (MOB Cosmetics, London, UK).

RESULTS AND DISCUSSION

Deflection and Stability Characterization under Radial Field Sweep

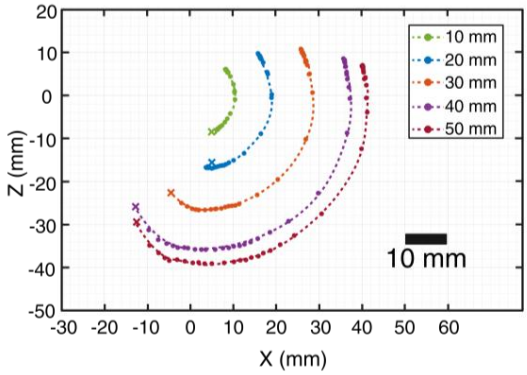


Figure 2. Comparison of planar quasistatic tip trajectories under rotating radial magnetic field conditions (0°-360°) for all deployment lengths (10-50 mm), each marked with an'x' to highlight the final stable tip position, before non-planar instability was observed.

Under high-angle fields (>180°), combined resistive torques limited maximum deflection to 240° and 250° fields for 10mm and 20mm deployment lengths, respectively. This produced a symmetrical, fan-like planar trajectory. Non-planar instability was constrained while elastic restoring torques promoted stable upward translation near a 360° field. At 30mm, transitional behaviour largely replicated the general fan trajectory, with stable retroflexion to a 290° field and minor midpoint sagging between 40°–70° fields. Magnetic torque generally balanced gravitational forces, enabling levitation and relatively linear deflection. At longer deployment lengths, geometrical anchoring concentrated mass, stiffness, and restoring forces proximally, maintaining a stable base while allowing for distal retroflexion up to 300° fields in free space.

In Vivo Localized Delivery over Simulated Skip Lesion

The MSCR's axial magnetization profile, combined with submillimeter-scale tip geometry allowed for access and optimal μ-channel placement over the target lesion (**Figure 3(B)**). By 15s, approximately 60% of the total visible lesion surface was covered with the drug simulant. By the end of the 30-second delivery period, a coverage of 86.1% was achieved with respect to the visible lesion area and 74.50% with respect to the total lesion surface area.

CONCLUSION

We highlight the clinical potential of an existing conical tapered MSCR design, which offers stable, high-angle deflections across various deployment lengths, and facilitates the direct integration of a μ-channel network for targeted fluidic delivery within a live porcine model. We demonstrate the design's capabilities to traverse dynamic, GI mucosal surfaces for optimal localized delivery across a millimeter-scale target lesion area.

REFERENCES

[1] Shah RS *et al*. Medical therapies for postoperative Crohn's disease. Ther Adv Gastroenterol. 2021;15(14).

[2] Kumar A *et al*. A review of the therapeutic management of Crohn's disease. Ther Adv Gastroenterol. 2022;17(15).

[3] Bacchetti A *et al*. Enhancing the functionality of soft continuum robots for minimally invasive and endoluminal interventions: a review. Prog Biomed Eng. 2026;8(2).

[4] Mao L et al. Magnetic steering continuum robot for transluminal procedures with programmable shape and functionalities Nat. Commun. 2026;15(3759).

[5] Bacchetti A *et al*. Tapered Magnetic Soft Continuum Catheters with Integrated Microchannels for Cerebral Intra-Arterial Chemotherapy Delivery. Soft Robotics, 2026; 0(0).

Forward-Model-Based Anomaly Detection for Magnetic Localization Systems

İ. Şahin[1], C. Öztürk[1,2], and O. Akanyeti[3]

[1]*Institute of Biomedical Engineering, Boğaziçi University, Türkiye*
[2]*Boğaziçi University Center for Targeted Therapy Technologies (Boğaziçi CT[3]), Türkiye*
[3]*Department of Computer Science, Aberystwyth University, UK*
ismail.sahin2@std.bogazici.edu.tr

INTRODUCTION

Magnetic localization systems based on Hall-effect sensor arrays have been widely studied for decades, particularly for clinical applications such as catheter tracking, capsule endoscopy and surgical navigation. To date, most research has focused on developing computationally efficient inverse models for predicting magnet pose from sensor measurements, ranging from physics-based dipole formulations to data-driven machine learning approaches [1], [2], [3]. In contrast, comparatively little attention has been given to ensuring robust performance of these models under adverse conditions, such as sensor malfunctions, temporary signal corruption, or external magnetic disturbances [1], [4], [5]. These conditions commonly arise in surgical environments, where sensors may fail intermittently and background electromagnetic noise from nearby electronic devices can introduce additional measurement variability.

To address this problem, we propose an anomaly detection architecture based on a forward model that predicts sensor responses from known magnet and sensor poses and enables direct comparison between expected and measured signals. The suitability of this forward-model-based approach for detecting anomalous behavior in magnetic localization systems is evaluated using experimental data collected under controlled laboratory conditions.

MATERIALS AND METHODS

A. Abnormal Event Detection and Classification

The architecture is inspired by predictive sensory processing in animals, in which expected sensor readings generated by an internal forward model are compared with real sensory measurements to produce a residual error signal [6]. The residual signal $r_i(t)$ for sensor i at time t is defined as the Euclidean distance between the measured and predicted three-channel magnetic field vectors, $\mathbf{B}_i(t)$ and $\hat{\mathbf{B}}_i(t)$, respectively.

External magnetic interference is assumed to influence multiple sensors simultaneously, whereas a faulty sensor typically exhibits abnormal behavior independently of neighboring sensors. Based on these assumptions, the number of sensors with elevated residuals, their spatial proximity, and the spatiotemporal patterns of these residuals are analyzed to distinguish sensor faults from external magnetic interference.

Using these criteria, events were broadly classified into four categories: (i) single-sensor fault (abnormal response isolated to a single sensor), (ii) multi-sensor fault (abnormal responses observed across multiple sensors without coherent spatial or temporal correlation), (iii) external interference (abnormal responses observed in a spatially clustered group of sensors, characterized by a clear peak in residual magnitude that decreases across neighboring sensors), and (iv) normal operating conditions (time windows in which no sensors exhibit elevated residuals). A sensor is considered to exhibit an abnormal response if its residual magnitude exceeds a predefined threshold T_i for a minimum duration τ, defined as $r_i(t) > T_i \quad \text{for} \quad t \in [t_0,\, t_0 + \tau]$.

Spatial proximity among sensors with abnormal responses was quantified using the mean pairwise Euclidean distance. To assess spatial magnitude patterns, the correlation between residual magnitude and distance from the peak active sensor was computed for each cluster, defining a spatial decay score. External magnetic interference was expected to produce negative decay values, indicating a decrease in residual magnitude with increasing distance from the peak sensor.

The forward model was implemented using a magnetic dipole formulation,

$$\hat{\mathbf{B}}_i = \frac{\mu_0}{4\pi}\left(\frac{3\left(\mathbf{m}\cdot\mathbf{r}_{\mathrm{rel},i}\right)\mathbf{r}_{\mathrm{rel},i}}{\|\mathbf{r}_{\mathrm{rel},i}\|^5} - \frac{\mathbf{m}}{\|\mathbf{r}_{\mathrm{rel},i}\|^3}\right), \qquad (1)$$

where $\mathbf{r}_{\mathrm{rel},i}(t) = \mathbf{r}_i - \mathbf{r}_0(t)$, $\mathbf{r}_i$ denotes the fixed position of sensor i, $\mathbf{r}_0(t)$ denotes the dipole (magnet) center position, and $\mathbf{m}(t)$ represents the magnetic dipole moment.

Assuming that changes in magnet pose between consecutive time steps are sufficiently small, ground-truth pose data at time step $(t-1)$ were used to predict $\hat{\mathbf{B}_i}(\mathbf{t})$ in order to evaluate the forward model independently of errors introduced by inverse-model estimation. In a complete real-time implementation, the magnet pose at $(t-1)$ can be estimated using an inverse model, such as the approach described in [7].

B. Setup and Experiments

Experiments were conducted in the Intelligent Systems Laboratory, Department of Computer Science, Aberystwyth University. The sensing board consisted of 64 Hall-effect sensors arranged in an 8×8 grid with 30 mm center-to-center spacing. Each sensor measured three-axis magnetic field components along the x, y, and z directions. A 6-DOF Universal Robots UR3e manipulator was used to move the target magnet along two predefined trajectories above the sensor array. Sensor readings and magnet poses were

recorded at 1 Hz. The target magnet was a cylindrical N52 neodymium magnet (D=3 mm and L=6 mm). Further details of the experimental setup are also provided in [7].

For each trajectory, noise-free data were first collected (1705 samples). This dataset was subsequently used to emulate sensor faults through offline data manipulation. Gaussian noise with a large standard deviation was added to the readings of randomly selected sensors across the array: i) one sensor was corrupted at a time to emulate a single-sensor fault, and ii) two sensors were corrupted simultaneously to emulate a multi-sensor fault.

The data collection trials were then repeated while a second cylindrical magnet (N52 neodymium magnet, D=6 mm and L=6 mm) was manually moved toward and away from the sensor board from different positions and approach angles. This second magnet acted as a dynamic interference source while the target magnet continued to follow its nominal trajectory. An example illustrating the field strength measurements during external magnetic interference is shown in Fig. 1.

C. Data Analysis

For each trajectory, 16 external magnetic interference events, 16 single-sensor fault events, and 16 multi-sensor fault events were generated. The resulting time-series data were processed using the proposed architecture, and classification performance was evaluated using precision, recall, and F1-score metrics. During data analysis, threshold parameters were fixed at $T = 200$ and $\tau = 5$ s.

I. CONCLUSIONS AND DISCUSSION

Table I summarizes the classification performance of the proposed architecture across the three anomaly classes. High precision and recall were achieved for all classes, with F1-scores consistently above 90%. In the noise-free dataset, the system generated only two false positives, one single-sensor fault and one multi-sensor fault, indicating reliable performance under normal operating conditions. Further analysis showed that these misclassifications were primarily associated with small forward-model prediction errors when the target magnet was close to the sensor board.

TABLE I: Anomaly classification performance.

Class	Precision (%)	Recall (%)	F1-score (%)
Single-sensor fault	94.1	100.0	96.9
Multi-sensor fault	84.0	98.4	90.5
External interference	93.7	93.7	93.5

The performance of the proposed anomaly detection architecture is promising; however, several factors influencing system reliability were identified. These include the accuracy of the forward and inverse models, the selection of amplitude and temporal thresholds for sensor activation, and the approach used to quantify spatial grouping of active sensors. A systematic investigation of these parameters is part of ongoing work within our research group.

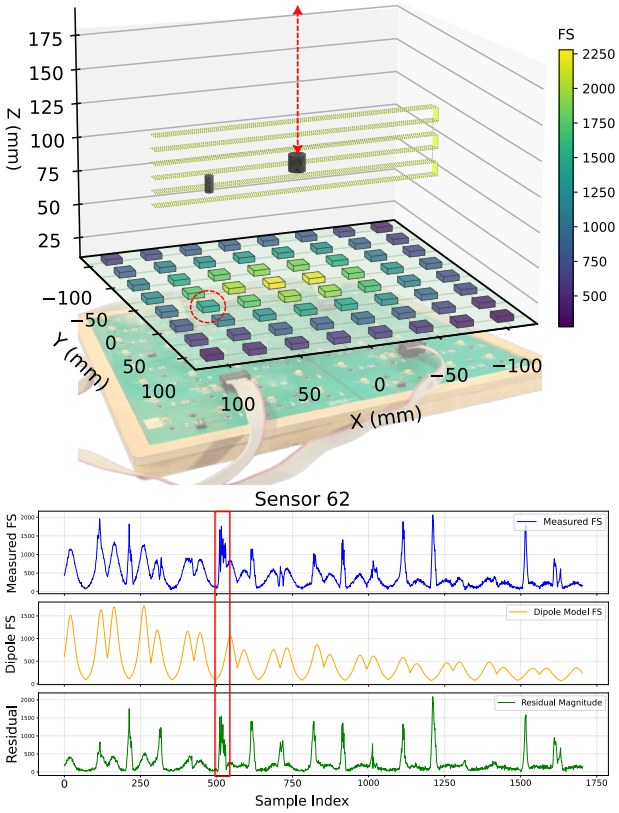


Fig. 1: Upper: Snapshot of measured field strength when both the target (smaller) and a second magnet (bigger) were present (yellow line: trajectory of the target magnet, red dashed line: direction from which the second magnet was introduced). Sensor 62 was marked with red dashed circle. The sensor board was also shown under the field strength measurements as a reference. Lower: Measured field strength, dipole-model prediction field strength, and residual signal for Sensor 62 (red square: the time point shown in the upper panel).

REFERENCES

[1] D. Son, S. Yim, M. Sitti, "A 5-D Localization Method for a Magnetically Manipulated Untethered Robot Using a 2-D Array of Hall-Effect Sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 2, pp. 708–716. 2016.

[2] M. Wu, T. Langerak, O. Hilliges, J. Zarate, "Using Synthetic Data in Supervised Learning for Robust 5-DoF Magnetic Marker Localization," *IEEE Transactions on Magnetics*, vol. 60, no. 8, Art. no. 4000211. 2024.

[3] Y. Wang, G. Pittiglio, X. Lin, P. E. Dupont, "Few-Shot Magnet Localization Using Sim-to-Real Transfer Learning," *IEEE Robotics and Automation Letters*, vol. 10, no. 12, pp. 12660–12667, 2025.

[4] S. Zeising, D. Anzai, A. Thalmayer, G. Fischer, J. Kirchner, "Innovative Differential Magnetic Localization Method for Capsule Endoscopy to Prevent Interference Caused by the Geomagnetic Field," *Advances in Radio Science*, vol. 19, pp. 207–213, 2021.

[5] E. A. Belalcazar-Bolaños, D. Torricelli, J. L. Pons, "Automatic Detection of Magnetic Disturbances in Magnetic Inertial Measurement Unit Sensors Based on Recurrent Neural Networks," *Sensors*, vol. 23, no. 24, p. 9683, 2023.

[6] D. Wolpert, M. Ghahramani, Z., M. I. Jordan, "An internal model for sensorimotor integration." *Science*, 269(5232), pp.1880-1882, 1995.

[7] İ. Şahin, C. Öztürk, O. Akanyeti, "Magnetic Localization for Tracking Cardiac Catheters in Time-Critical Surgical Operations." *In 2025 25th International Conference on Digital Signal Processing (DSP)*, pp. 1-5, 2025.

Magnetic–Fluidic Soft Robotic Catheter Tip for Enhanced Access in Superior Mesenteric Artery Aneurysm Intervention

Muhammad A. Khalid, Pietro Valdastri, and James H. Chandler

School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K.

INTRODUCTION

The superior mesenteric artery (SMA) arises from the abdominal aorta and gives rise to multiple branches that supply the pancreas, small intestine, and right colon. An SMA saccular aneurysm, although rare, is clinically significant and may develop in the main SMA trunk (approximately 6–8 mm in diameter) due to focal arterial wall degeneration combined with high-flow hemodynamic stress, particularly at branching points. The preferred management is an endovascular approach, especially for saccular lesions, typically involving coil embolization or thrombin-based embolization to exclude the aneurysm from circulation while preserving distal perfusion [1].

Traditional catheter-based systems face challenges in navigation and control, including dependence on operator skill, limited steerability, and reduced intraprocedural stability [2], which may lead to vascular injury, embolic agent malposition or migration, or aneurysm perforation. To address these limitations, medical continuum robots have been proposed using mechanical tendon/pull-wire, fluidic pneumatic/hydraulic, and magnetic actuation mechanisms [3]. Among these, magnetic actuation offers advantages in miniaturization and wireless control via external magnetic fields, while fluidic actuation can provide complementary functions such as variable stiffness, local elongation-contraction, and bending. However, active axial length control at the distal catheter tip remains limited.

In this work, we propose a dual-actuated soft continuum robotic catheter tip (SCR-CT) that combines magnetic steering with pressure- and vacuum-driven axial length control for enhanced access in SMA aneurysm intervention. Bidirectional length control increases the effective tip workspace during embolization, potentially reducing reliance on uncontrolled proximal manual catheter manipulation and minimizing vessel trauma. Magnetic steering provides three-dimensional (3D) maneuverability without internal transmission elements such as tendons, thereby reducing unwanted bending and improving navigational control. Fig. 1(a) shows the proposed application scenario, while Fig. 1(b) illustrates the catheter tip design and components. A kinematic model is developed and compared with experimental characterization under combined pneumatic and magnetic actuation.

MATERIALS AND METHODS

The fluidic element of the SCR-CT was based on a bellows actuator design and fabricated using 3D-printed injection molds and silicone rubber (Dragon Skin 20, Smooth-On, USA). The actuator was assembled with connectors and tubing to form a closed chamber. A ring-shaped N35 neodymium permanent magnet with an outer diameter of 4 mm, an inner diameter of 2 mm, and a length of 2 mm was attached to the distal tip. An Euler-Bernoulli beam theory-based model [4] was used to predict the tip position and bending angle at three pneumatically actuated lengths (20.5, 33.0, and 37.1 mm), using a Young's modulus of 338 kPa for Dragon Skin 20. The model approximates the bellows as an equivalent tube with averaged outer and inner diameters of 3.8 mm and 2.4 mm, respectively. The following two nonlinear equations were solved simultaneously using *fsolve* solver (MATLAB, MathWorks Inc., USA):

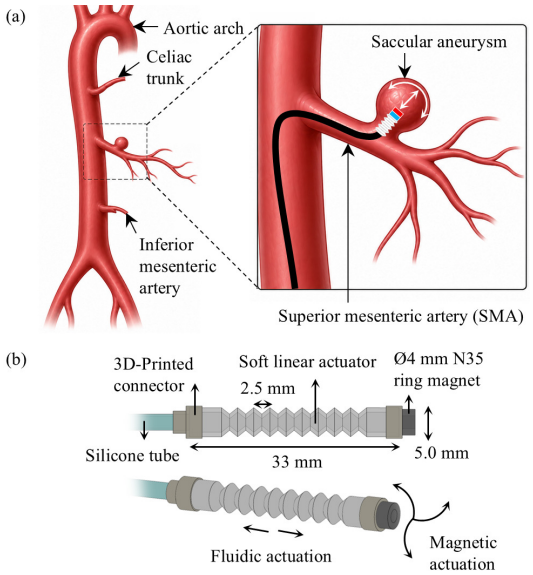


Fig. 1: (a) Magnetically steerable and pneumatically extendable robotic catheter tip positioned at the aneurysm neck, illustrating improved local access. (b) Design specifications and components of the catheter tip.

$$\tau = |\boldsymbol{M}||\boldsymbol{B}|sin(\varphi - \theta_L) \quad (1)$$

$$L = \frac{K\theta_L}{\tau} \quad (2)$$

where, $\boldsymbol{M}$ is the magnetic moment, $\boldsymbol{B}$ is the uniform magnetic field, τ is the magnetic torque, θ_L is the bending angle, φ is the applied magnetic field angle, L is the catheter length, and K is the bending stiffness. Pneumatic actuation was controlled using a stepper motor-driven syringe system, with chamber pressure monitored by a pressure sensor (2513254510491, Würth Elektronik, Germany). Magnetic

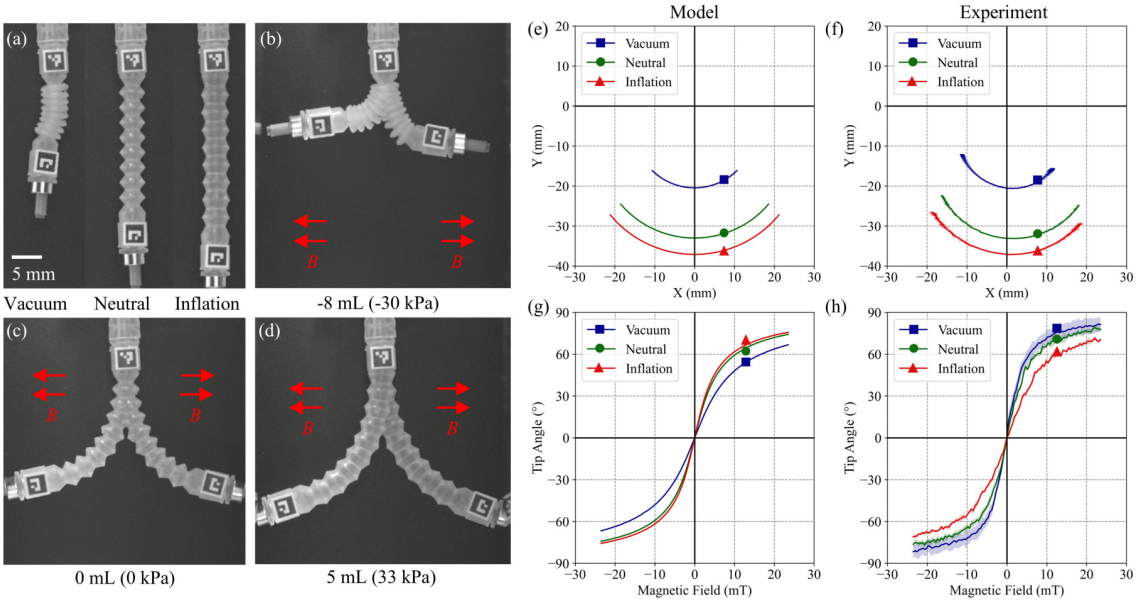


Fig. 2: (a) Pneumatic actuation at different input volumes (-8 mL: vacuum, 0 mL: neutral, and 5 mL: inflation). (b–d) Combined pneumatic and magnetic actuation up to ±23.5 mT. Model predicted and experimentally measured tip position ((e), (f)) and bending angle ((g), (h)). Model and experimental curves consist of 101 data points sampled at 0.47 mT increments; markers are used only for line identification.

actuation was performed using a DXHC10-200 Helmholtz coil system (Dexing Magnet Tech, China), which generated homogeneous magnetic fields up to ±23.5 mT at a field angle of 90°. Two ArUco markers placed at the catheter base and tip were tracked using a camera (Basler AG, Germany), and images were analyzed in Python/OpenCV to extract tip position and orientation.

RESULTS AND DISCUSSION

The pneumatic actuation, comprising inflation (5 mL) and vacuum (-8 mL), produced SCR-CT extension and contraction of 4.1 mm and 12.6 mm at 33 kPa and -30 kPa, respectively, resulting in a length variation of 16.7 mm, as shown in Fig. 2(a). Bidirectional length control enables compact deployment and increases the tip workspace through extension and retraction. The applied pressure range is lower than values reported for soft fluidic catheter systems [5]; however, future work will investigate saline-based fluidic actuation for improved intravascular safety. Magnetic actuation was performed at the three input volumes across a magnetic field range of ±23.5 mT, as shown in Fig. 2(b-d). The position and orientation plots from the model predictions and experiments (three repeats) are shown in Fig. 2(e, g) and Fig. 2(f, h), respectively. The predicted and measured tip position coordinates and bending angles show good agreement. The maximum bending angle was 81°, exceeding the typical aortomesenteric angle (approximately 30–70°) [6], suggesting sufficient steering capability for initial SMA engagement, with tortuous navigation to be evaluated in vascular phantoms. Root mean squared error (RMSE) values of 20°, 5°, and 7° were obtained for the bending angle, and 3.6 mm, 1.8 mm, and 3.4 mm for the tip position under vacuum, neutral, and inflation states, respectively. The RMSE was computed over the full range of applied magnetic fields.

These results support modeling the bellows as a tube with averaged dimensions. However, under vacuum, the bending angle exceeded the predicted value, as shown in Fig. 2(b). This discrepancy is attributed to state-dependent stiffness changes caused by actuator geometry, with bending stiffness decreasing under vacuum and increasing under inflation. Future work will account for this stiffness variation and contact modeling to improve the current free-space model. In conclusion, the proposed dual-actuated SCR-CT demonstrates controlled length variation and magnetic steering with potential to improve workspace access during SMA aneurysm intervention. Future work will include a working lumen for guidewire-compatible operation, anatomically realistic vascular phantom testing, variable magnetic-field orientations, and evaluation using saline-based fluidic actuation.

REFERENCES

[1] G. S. Pitcher, N. C. Cirillo-Penn, B. C. Mendes, F. Shuja, R. R. DeMartino, M. Kalra, T. C. Bower, W. S. Harmsen, and J. J. Colglazier, "Aneurysms of the superior mesenteric artery and its branches," *Journal of vascular surgery*, vol. 76, no. 1, pp. 149–157, 2022.

[2] T. da Veiga, J. H. Chandler, P. Lloyd, G. Pittiglio, N. J. Wilkinson, A. K. Hoshiar, R. A. Harris, and P. Valdastri, "Challenges of continuum robots in clinical context: a review," *Progress in Biomedical Engineering*, vol. 2, no. 3, p. 032003, 2020.

[3] Y. Peng, X. Liu, K. F. Chan, X. Song, and L. Zhang, "Robotic-assisted endovascular embolization: Progress and future perspectives," *SmartBot*, vol. 1, no. 1, p. e12009, 2025.

[4] M. A. Khalid, P. Valdastri, and J. H. Chandler, "3d-printed variable-length magnetically steered catheter tip for enhanced access in aneurysm interventions," in *2025 IEEE 8th International Conference on Soft Robotics (RoboSoft)*, 2025, pp. 1–7.

[5] T. Gopesh, J. H. Wen, D. Santiago-Dieppa, B. Yan, J. S. Pannell, A. Khalessi, A. Norbash, and J. Friend, "Soft robotic steerable microcatheter for the endovascular treatment of cerebral disorders," *Science robotics*, vol. 6, no. 57, p. eabf0601, 2021.

[6] A. Lurie, L. Guntupalli, V. Stoll, and A. Moon, "Superior mesenteric artery syndrome masquerading as irritable bowel syndrome: a case report," *Cureus*, vol. 15, no. 8, 2023.

A deployable, magnetically driven soft robotic platform for laparoscopic surgery: modeling, design and preliminary validation works

Matteo Bernabei[1], Lorenzo Mocellin[1], Matteo Cianchetti[1], Arianna Menciassi[1], Cesare Stefanini[2], and Gastone Ciuti[1]

[1]*The BioRobotics Institute, School of Advanced Studies Sant'Anna, Pontedera, Italy*
[2]*Department of Robotics, MBZUAI (Mohamed bin Zayed University of Artificial Intelligence), Abu Dhabi, UAE*
[1]email: matteo.bernabei@santannapisa.it

INTRODUCTION

In abdominal surgery, multi-port laparoscopy reduces invasiveness with respect to open surgery approach but requires multiple incisions and is affected by instrument collisions and motion constraints. Single-port surgery further limits access-related trauma, yet introduces challenges such as instrument crowding and reduced triangulation. Continuum and endoscopic systems for NOTES (Natural Orifice Translumenal Endoscopic Surgery) [1] exploit structural compliance to enhance access to confined anatomical regions; however, they face limitations in navigation and tissue interaction, and their clinical applicability remains strongly task and site-dependent. As novel devices continue to reduce both invasiveness and procedural trauma in state of the art procedures, magnetically actuated devices have been proposed to allow untethered external coupling of the internal laparoscopic tool [2]–[4]. In this context, magnetic coupling can serve also as anchoring means of laparoscopes [5] and actuation for graspers [6]. Here, the authors present a modular, deployable, self-assembling, magnetically driven soft robotic platform aimed at overcoming port crowding in single-port laparoscopy, validated through magnetic modeling, component fabrication, and preliminary in-vitro experiments.

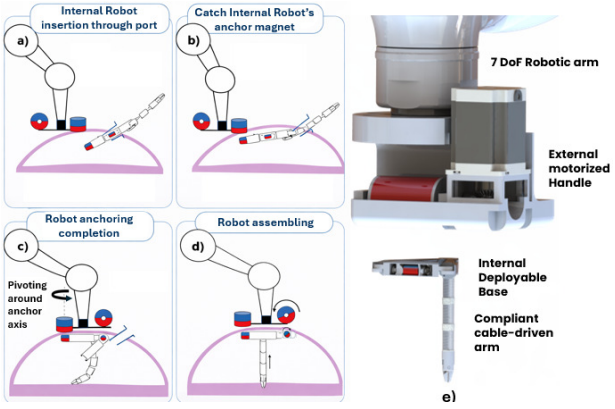


Fig. 1: Conceptual deployment sequence for the proposed single-port magnetically driven platform (a–d) and rendered overview of the complete dual system (e).

MATERIALS AND METHODS

A. Magnetic Modeling and Design of a Magnetic Clutch System for Cable-Driven Transmission

The system comprises two magnetically coupled modules enabling tetherless abdominal anchoring, planar translation, torque transmission for the soft arm's control. Fig. 1 illustrates the deployment of the magnetically driven system in a single-port laparoscopic scenario. The internal robot is introduced through the trocar with the base module in its insertable configuration, followed by the arm modules and gripper connected by actuation cables. After magnetic anchoring between internal and external magnets is established, diametral magnetic coupling enables actuation of the cable-driven system, allowing both assembly of the structure and subsequent arm bending and gripper closure. The permanent magnet configuration embedded in both modules was analyzed in a previously validated, custom built in-silico environment developed in MATLAB (The MathWorks, Natick, MA, USA), where magnets were modeled using the magnetic charge formulation [7] and represented through equivalent volumetric magnetic charge densities. Fig. 2 b-c) illustrate the simulated interaction forces. During deployment, the internal platform self-assembles and actuates a modular articulated arm via a miniaturized cable-driven mechanism powered by the integrated magnet (Fig. 2 a)). An externally driven rotation of the diametrically magnetized internal magnetic pulley spools two 0.25 mm nylon cables, while a 5° pivoting motion engages a bevel gear–based rotation switch that enables switching between synergistic and antagonistic cable-pulling modes for assembly, grasper closure, and controlled arm bending. The mechanism components were fabricated in bio-compatible M3 Crystal resin via SLA 3D printing, while the magnetic pulley shaft was manufactured in brass.

B. Deployable Modular Structure and Soft Arm Design

The mechanical design of the deployable modular platform is presented, focusing on the internal device's triangle-shaped base and the integrated soft robotic arm. Fig. 3a) shows the fully integrated laparoscopic device concept within undergoing chracterization tests for bending tracking (Fig. 3b). Fig. 3c) shows the robot's base module when in its insertable configuration. The lenght of the three segments of the triangle is 70 mm and their cross-section

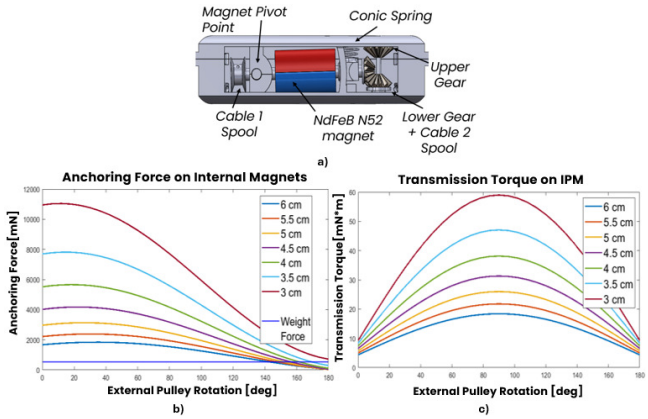


Fig. 2: (a) The design of the magnetic clutch mechanism.(b) Estimated anchoring force on the internal magnets under phase shift between the magnetic pulleys at different operative distances. (c) Estimated torque transmitted to the internal pulley magnet (IPM) under the same conditions.

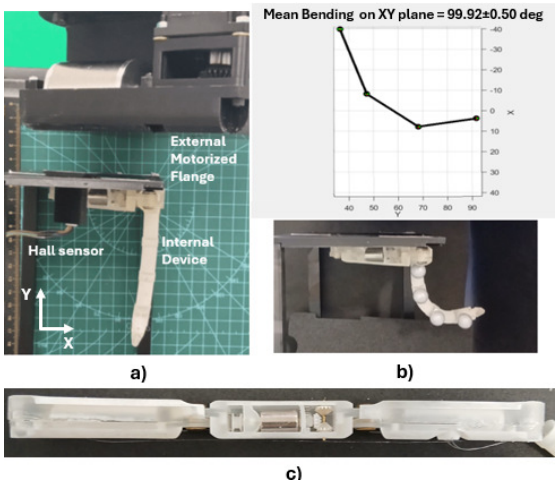


Fig. 3: (a) The assembled dual robotic system in optical tracking setup to assess soft arm's bending. (b) Mean maximum arm bending angle reached (larger than 90°). (c) Assembled deployable base module, shown in insertable (open) configuration.

measures 14.5 mm. The soft arm consists of three spring-shaped, cable-driven modules and a laparoscopic grasper at the distal tip. Each module comprises a helical elastic body with complementary conical interfaces enabling axial alignment during deployment, and three columns of cylindrical rolling contacts that allow controlled bending under antagonistic cable actuation. A parametric model derived from [8], in which the bending angle per coil is expressed as $\theta = n\gamma\pi M r_{mc}\left(\frac{1}{EI} + \frac{1}{GJ}\right)$ was employed to optimize the module geometry across multiple parameter configurations. The resulting maximum bending of 90° across the three-module assembly, with each module contributing up to 30° of deflection, was validated through bending tracking and tensile tests on a dedicated Universal Testing Machine.The functional capability of the assembled platform was assessed through an ex-vivo resection task, demonstrating tissue handoff, lift-and-cut, and sample lift-and-hold operations, as reported in Fig. 4.

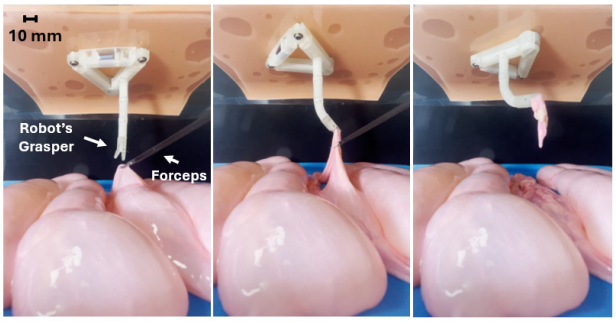


Fig. 4: Resection task main sequence on ex-vivo tissue.

CONCLUSIONS AND DISCUSSION

This study presents the development and preliminary validation of a fully untethered, magnetically driven robotic platform designed through a simulation-based approach. Experiments demonstrated the feasibility of single-port laparoscopic deployment, stable magnetic anchoring, and effective internal actuation, including rotational and translational motion and soft-arm bending up to 90°. Future work will address quantitative evaluation under surgical-like conditions and iterative design optimization, targeting material selection for components subjected to magnetic interaction forces, with requirements of mechanical durability, biocompatibility, and low friction, and the transition from the current rigid three-link base to a compliant deployment structure.

Acknowledgements

We acknowledge the support of the European Union through the Next Generation EU project ECS00000017, *'Ecosistema dell'Innovazione' Tuscany Health Ecosystem (THE, PNRR, Spoke 9: Robotics and Automation for Health)'*.

REFERENCES

[1] H. Wu, D. Xue, M. Deng, and et al., "Progress, challenges, and future perspectives of robot-assisted natural orifice specimen extraction surgery for colorectal cancer: a review," *BMC Surgery*, vol. 24, p. 255, 2024.

[2] L. Magnetics, "Magnetic surgery: Levita magnetic surgical system," Patent US 11,020,137, 2021, fDA DEN150007 Summary.

[3] S. Tognarelli *et al.*, "Magnetic anchoring and guidance systems for notes procedures," *Surgical Endoscopy*, vol. 26, no. 10, pp. 2867–2871, 2012.

[4] N. Garbin *et al.*, "Magnetically driven robotic device for minimally invasive surgery," *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1358–1371, 2015.

[5] J. Li, T. Zhang, T. Cheng, Y. Li, C. Ng, P. Chiu, and Z. Li, "Design, analysis, and preliminary validation of magnetic anchored and cable driven endoscope for minimally invasive surgery," *IEEE Transactions on Medical Robotics and Bionics*, vol. PP, pp. 1–1, 11 2024.

[6] C. He, R. Nguyen, H. Mayer, L. Cheng, P. Kang, D. A. Aubeeluck, G. Thiong'o, E. Fredin, J. Drake, T. Looi, and E. Diller, "Magnetically actuated dexterous tools for minimally invasive operation inside the brain," *Science Robotics*, vol. 10, no. 100, March 2025.

[7] E. P. Furlani, *Permanent Magnet and Electromechanical Devices*, ser. Electromagnetism, 2001.

[8] Y. Hu, L. Zhang, W. Li, and G.-Z. Yang, "Design and fabrication of a 3d printed metallic flexible joint for snake-like surgical robot," *IEEE Robotics and Automation Letters*, vol. PP, pp. 1–1, 01 2019.

AI Generated Synthetic Data for Accelerating the Development of Data-Driven Magnetic Localization Models

İ. Şahin[1], C. Öztürk[1,2], and O. Akanyeti[3]

[1]*Institute of Biomedical Engineering, Boğaziçi University, Türkiye*
[2]*Boğaziçi University Center for Targeted Therapy Technologies (Boğaziçi CT[3]), Türkiye*
[3]*Department of Computer Science, Aberystwyth University, UK*
ismail.sahin2@std.bogazici.edu.tr

INTRODUCTION

In medicine, recent advances in machine learning, particularly deep learning, have led to growing interest in data-driven approaches for magnetic pose estimation to track in-body devices such as catheters and capsule endoscopes [1], [2], [3], [4], [5]. These methods often achieve high accuracy in controlled laboratory environments. However, they are prone to overfitting when training data are limited. Acquiring comprehensive datasets with known magnet poses and corresponding sensor measurements is technically challenging, as collecting measurements across the full range of possible positions and orientations is often impractical.

One potential strategy to address this limitation is the generation of synthetic training data. If realistic magnetic field responses can be generated computationally for arbitrary magnet poses and sensor locations, synthetic datasets could complement experimental measurements and increase training diversity, thereby improving the robustness of data-driven localization models. Traditionally, synthetic magnetic field data have been generated using physics-informed models derived from analytical formulations of magnetic dipole fields [4], [5]. While such models are theoretically grounded, their practical performance is often limited by deviations between idealized assumptions and real-world conditions [3], [4].

In this study, we investigate whether an AI-based model can be trained to generate synthetic magnetic field measurements that better reflect real sensor responses than conventional physics-informed models. Such a capability has the potential to accelerate the development of data-driven magnetic localization systems and improve their generalization to previously unseen configurations.

MATERIALS AND METHODS

Experimental setup and data acquisition

Experiments were conducted in the Intelligent Systems Laboratory, Department of Computer Science, Aberystwyth University. The sensor array consisted of 64 three-axis sensors arranged in an 8×8 grid with 30 mm spacing. The physical footprint of the sensor array was $210\,\text{mm} \times 210\,\text{mm}$ in the x-y plane, and the z-direction was defined as perpendicular to the sensor array surface. A 6-DOF Universal Robots UR3e manipulator was used to move

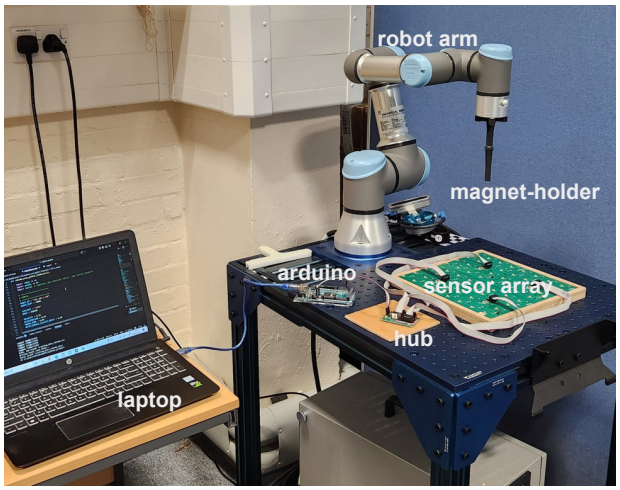


Fig. 1: Experimental setup and magnetic sensor array.

a cylindrical N52 neodymium magnet (with a diameter of 3 mm and a length of 6 mm) along a predefined trajectory covering a volume defined by $x, y \in [-105, 105]$ mm and $z \in [40, 172]$ mm. The magnet orientation was varied within $[-180, 180]^\circ$ in azimuth and $[-2, 90]^\circ$ in elevation.

Data were acquired at a sampling rate of 1 Hz and included magnetic field measurements from all sensors together with the corresponding magnet position and orientation. Overall, the dataset comprised 230,391 samples.

Generating synthetic data using AI models

Model architecture: The synthesis model was implemented as a residual multilayer perceptron (ResMLP) consisting of a linear embedding layer, followed by six residual MLP blocks. Each block comprised LayerNorm, two linear layers with a GELU activation in between, and an identity residual connection. The model used a feature dimension of 128 and an expansion ratio of 4. A final MLP head mapped the resulting representation to the magnetic field outputs of the held-out sensor group.

Training approach: From the total set of 64 sensors, eight datasets were generated using a leave-sensor-out strategy. In each dataset, measurements from 56 sensors were used for training, while data from the remaining eight sensors were held out for testing. This process resulted in eight distinct training-testing splits, each corresponding to a

different subset of held-out sensors. Eight separate ResMLP models were trained using a batch size of 2048 for 30 epochs with the AdamW optimizer.

Generating synthetic data using dipole model

The proposed AI-based synthetic data generation approach was compared against a conventional physics-based dipole model. Approximating the magnet as an ideal magnetic dipole, the magnetic field at sensor i at time step t was computed as

$$\mathbf{B}_i(t) = \frac{\mu_0}{4\pi}\left(\frac{3(\mathbf{m}\cdot\mathbf{r}_{\mathrm{rel},i})\mathbf{r}_{\mathrm{rel},i}}{\|\mathbf{r}_{\mathrm{rel},i}\|^5} - \frac{\mathbf{m}}{\|\mathbf{r}_{\mathrm{rel},i}\|^3}\right), \qquad (1)$$

where $\mathbf{r}_i$ denotes the position of sensor i, $\mathbf{r}_0$ denotes the dipole (magnet) center position, $\mathbf{r}_{\mathrm{rel},i} = \mathbf{r}_i - \mathbf{r}_0$ is the relative position vector between the magnet and sensor i, and $\mathbf{m}$ represents the magnetic dipole moment.

Performance evaluation

Measured and predicted magnetic field responses were first compared at the sensor level using three complementary similarity metrics: Pearson correlation coefficient, root mean square error, and cosine similarity.

The impact of magnetic field prediction errors was then evaluated in the context of downstream magnet pose estimation. While similarity metrics quantify agreement at the sensor level, localization performance provides a more application-relevant assessment of whether the generated synthetic data are sufficiently accurate for practical use.

To perform this evaluation, a localization model was trained exclusively using real measured magnetic field data. The model received magnetic measurements from all sensors together with their spatial coordinates as input. The localization network was implemented using a one-dimensional residual convolutional network architecture with a feature dimension of 64 and a depth of 4 residual blocks. The network took sensor readings and sensor positions as a flattened one-dimensional array and estimated the magnet's position and orientation.

During inference, predicted magnetic field measurements generated by either the AI-based model or the dipole model were used as inputs to the trained localization model to determine whether these synthetic measurements were sufficient to recover accurate magnet pose estimates. For the AI-based approach, predictions from the eight ResMLP models corresponding to the held-out sensor subsets were combined to reconstruct a complete set of synthetic measurements equivalent to those from a full 64-sensor array.

Pose estimation accuracy for dipole and AI-based predictions was quantified using the Euclidean distance between predicted and reference magnet position and orientation vectors. Statistical significance between dipole-based, AI-based, and real measurement-based predictions (used as baseline) was assessed using the Wilcoxon signed-rank test with a significance level of $p < 0.05$.

TABLE I: Comparison between predicted and measured sensor readings.

Metric	Dipole	AI	p-value
Pearson corr. coeff.	0.994 ± 0.001	0.997 ± 0.001	$p < 0.05$
RMSE (LSB/μT)	73.7 ± 8.2	49.5 ± 7.1	$p < 0.05$
Cosine similarity	0.658 ± 0.12	0.74 ± 0.10	$p < 0.05$

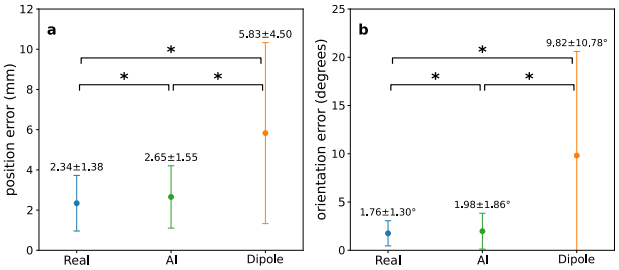


Fig. 2: Comparison of real, AI-generated, and dipole-generated data during magnet pose estimation: (**a**) position error and (**b**) orientation error. ($*$: $p < 0.001$).

CONCLUSIONS AND DISCUSSION

At the sensor level, AI-generated sensor readings were consistently closer to measured sensor readings than dipole-generated readings across all performance metrics (Table I).

The improvements observed at the signal level also translated into improved downstream localization performance (Fig. 2). Pose estimation using AI-generated synthetic data was substantially more accurate than localization based on dipole-generated synthetic data and approached the performance achieved using real sensor measurements, although the difference remained statistically significant.

Our results suggest that synthetic data generated by the AI-models have the potential to enhance the development of data-driven magnetic localization models by improving the quality and diversity of training data. However, further evaluation using independently collected test datasets is required to corroborate these results, as data-driven models are known to suffer from overfitting when assessed only on held-out data from the same acquisition session. Such independent validation remains an important part of ongoing research within our group.

REFERENCES

[1] D. Son, S. Yim, and M. Sitti, "A 5-D Localization Method for a Magnetically Manipulated Untethered Robot Using a 2-D Array of Hall-Effect Sensors," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 2, pp. 708–716, Apr. 2016.

[2] Y. Wang, G. Pittiglio, X. Lin, and P. E. Dupont, "Few-Shot Magnet Localization Using Sim-to-Real Transfer Learning," *IEEE Robotics and Automation Letters*, vol. 10, no. 12, pp. 12660–12667, Dec. 2025.

[3] B. Lv, H. Dai, Y. Qin, and H. Lin, "A High-Precision Magnetic Localization Method Based on EFLM for Attenuating Interference," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, Art. no. 6504310, 2025.

[4] M. Wu, T. Langerak, O. Hilliges, and J. Zarate, "Using Synthetic Data in Supervised Learning for Robust 5-DoF Magnetic Marker Localization," *IEEE Transactions on Magnetics*, vol. 60, no. 8, pp. 1–11, Aug. 2024.

[5] M. R. Zarifi Khameneh, M. Jamali, and H. N. Pishkenari, "Symmetry-Based Independent Fully Convolutional Neural Networks for Magnetic Localization Using Sensor Array," *IEEE Sensors Journal*, vol. 25, no. 19, pp. 36873–36880, Oct. 2025.

Bi-directional Magnetic Vine Robot for Enhanced Endoscopic Navigation

Alexander Brown[1], Alistair Bacchetti[1], Benjamin Calmé[1], Joshua Davy[1], Sukjun Kim[2], Tania K. Morimoto[2], James H. Chandler[1], and Pietro Valdastri[1]

[1]*School of Electronic and Electrical Engineering, University of Leeds*
[2]*Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92093, USA*

INTRODUCTION

Vine robots have shown significant promise in the advancement of endoluminal surgery procedures through their ability to navigate complex and tortuous anatomies [1]. Their motion is driven by internal pressure, which everts material at the distal end to extend the robot body and enables shear-free locomotion [2]. Retraction involves inverting this material to reduce the deployed length. However, existing approaches such as tail pulling can induce buckling, where the robot body kinks laterally, preventing further inversion and undermining the advantages of shear-free movement [3].

Magnetic actuation offers wireless torque transmission without embedded electronics at the tip, and recent work has demonstrated magnetic steering of vine robots [4] and linear retraction [5]. However, magnetically aided retraction through curved environments with bi-directional path selection capability has not yet been realized.

This work presents a Magnetic Vine Robot (MVR) capable of multi-path exploration. Through wireless torque transmission from a rotating External Permanent Magnet (EPM) to a diametrically magnetized Internal Permanent Magnet (IPM) housed within internal, tip-mounted rollers, we demonstrate bi-directional navigation (repeated directional advancement and curvilinear retraction) as illustrated in Fig. 1. The resulting 25.5 mm diameter MVR achieves sequential navigation and retraction through five branches offset by ± 30° over a 120° range, also enabling integrated tool channel deployment through the MVR tail, demonstrating the potential of the system for use in endoscopy procedures.

MATERIALS AND METHODS

System Design and Fabrication

The MVR (Fig. 2) comprises four integrated components: (i) Internal Magnetic Capsule (IMC) containing a 6 mm diameter, 20 mm long tip rollers, one passive and one housing a 5 mm diameter, 10 mm long diametrically magnetized IPM; (ii) 20 mm diameter, 20 mm long diametrically magnetized rotating EPM; (iii) Motorized 10 mm diameter base rollers, used for MVR tail conditioning; (iv) 25.5 mm diameter vine robot body fabricated from 38 μm thick thermoplastic-elastomer film (Stretchlon 200, Airtech Advanced Materials, UK) using laser welding. The IMC and rollers were 3D printed in clear resin (Form 4, FormLabs, USA) with knurled surface features for traction on the MVR tail. The IPM (Tube0140D, SuperMagnetMan, USA), was

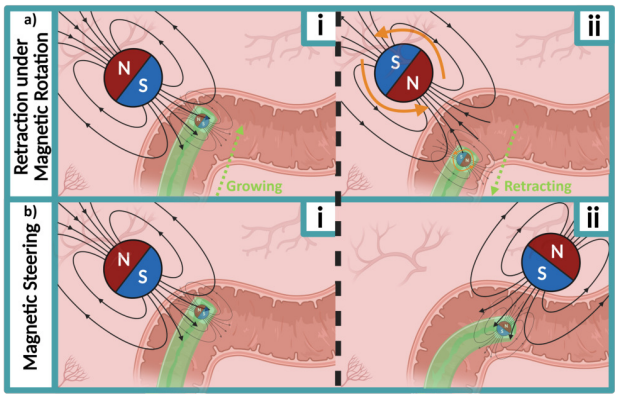


Fig. 1: Overview of the MVR. (a) Magnetic coupling between rotating EPM and IPM during (i) growth and (ii) retraction. (b) Magnetic steering: (i) controlled growth to a point of curvature. (ii) EPM steers the MVR tip to navigate the environment.

press-fit into the active roller, with integrated bearings in the IMC to support the roller shafts.

During deployment, the EPM is magnetically coupled to the IPM, driving the tip rollers when rotated and inducing steering when translated relative to the tip. The EPM rotation is synchronized with the base rollers such that as tail material is released from the base, the tip rollers simultaneously evert previously-fed tail material at the distal end, maintaining nominal tension in the system to avoid poor tail conditioning at the tip rollers, and allow for continuous forward growth. Unsynchronized rotation results in tail slack and roller jamming. Internal pressure maintains the cylindrical structure of the MVR (allowing for translation of the IMC through the body) and assists eversion. Retraction reverses this process; the EPM drives the tip rollers in the opposite direction to draw the tail inward and invert material at the tip, while the base rollers maintain tension preventing roller jamming. The magnetic coupling enables wireless torque transmission, where the maximum torque generated (and the resultant axial force transmitted by the tip rollers onto the tail) must be sufficient to overcome the resistive retraction forces from the growing pressure, material yield, and distributed friction.

Experimental Validation

Multi-Branch Path Selection: To evaluate the bi-directional navigation capability, the MVR was grown sequentially into each of five branching channels with fixed bend angles of 0° (centerline), +30°, +60°, -30° and -

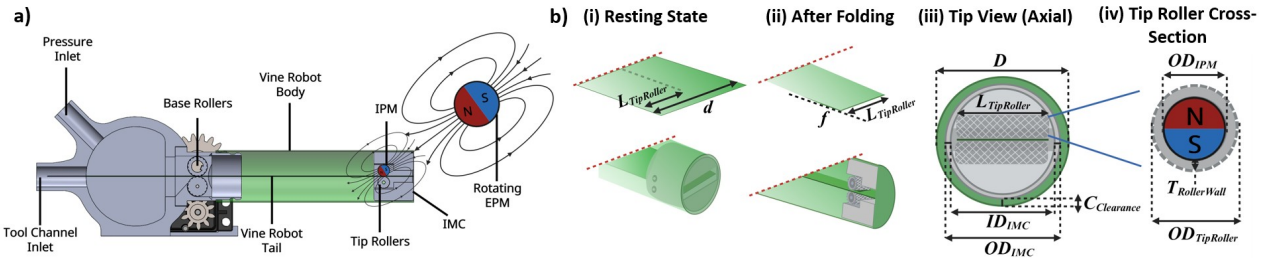


Fig. 2: MVR System Design. (a) Cross-sectional view showing the air tank with ports to allow pressure in and tool channel where the tail is anchored, enabling medical tooling to be passed through to the tip. Motorized base rollers for tail conditioning. Vine robot consisting of the tail and body. IMC at the tip of the vine housing the tip rollers and IPM. Rotating EPM magnetically coupled with the IPM. (b) Schematics of the vine robot and IMC: (i) Vine robot in resting state, (ii) Vine robot after folding for miniaturization and alignment with the tip rollers, (iii) Cross-sectional view of the IMC, (iv) Cross-sectional view of active tip roller, showing IPM integration.

60°. Following complete deployment into each channel, the EPM was actuated in reverse to actively retract the MVR from the channel. After complete retraction from each branch, the base rollers were locked in position and EPM was repositioned to align the tip of the MVR with an alternative branch angle, where the MVR was successfully redeployed, demonstrating bi-directional path selection capability. Successful advancement and retraction was achieved across all tested angles without buckling or loss of magnetic coupling (Fig. 3a). The effective tip roller force was measured as 1.42 N at a 25 mm EPM-IPM separation distance through linear retraction experiments, with the minimum applied pressure ranging between 2.6 kPa and 4.5 kPa.

Tool Channel Integration: A 1.75 mm diameter lumen was integrated through the center of the passive base and tip rollers by creating a 'V-shaped' groove. To demonstrate surgical tool deployment capability, 2.4 FR diameter grasping forceps (5720001 Surlok Stone Grasping Forceps, Gyrus Acmi, Japan) were inserted through the tool channel in the air tank where the tail is anchored. The EPM position and rotation were adjusted to manipulate the MVR tip position, enabling the forceps to successfully grasp and retrieve a target object (Fig. 3b). This demonstrates the potential of the MVR for use in surgical procedures with instrument delivery and task execution.

CONCLUSIONS AND DISCUSSION

This work addresses the challenge of retraction in vine robots by implementing magnetically actuated tip rollers for bi-directional path exploration. Through wireless torque transmission between the EPM and IPM embedded within the active tip roller, the 25.5 mm diameter MVR achieved successful retraction across tested branch angles (± 30°, ± 60°) without buckling. The bi-directional navigation capability enables sequential path selection, allowing the MVR to retract from one branch and redeploy into alternative pathways. This capability has the potential to allow for multiple task or navigation attempts during endoscopic procedures, improving access to challenging anatomical targets

Current limitations include the size of the EPM that restricts the EPM-IPM separation distance to a maximum

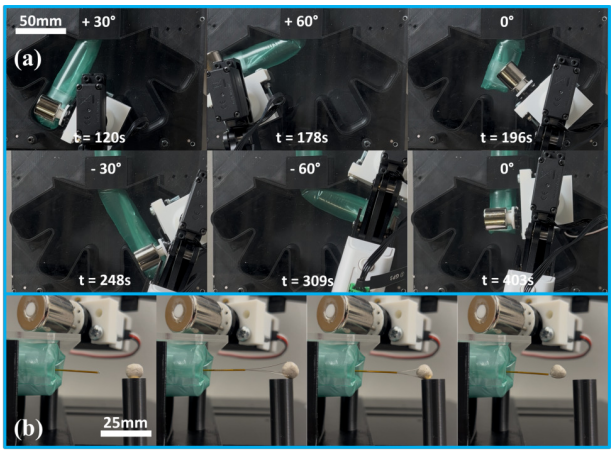


Fig. 3: Bi-directional navigation and tool channel integration. (a) Multi-branch path selection showing sequential steering, navigation and retraction through channels at angles from -60° to +60°, demonstrating bi-directional exploration capability. (b) Tool channel deployment with grasping forceps retrieving target object, showing instrument delivery and task execution capability.

of 40 mm, and the lack of integrated visual feedback at the tip. Future work will address camera integration within the IMC to allow for navigation without operator line of sight inside an anatomically accurate colon phantom.

REFERENCES

[1] S. Al Harthy, S. H. Sadati, C. Girerd, S. Kim, A. Mondini, Z. Wu, B. Saldarriaga, C. A. Seneci, B. Mazzolai, T. K. Morimoto *et al.*, "Tip-growing robots: Design, theory, application," *IEEE Transactions on Robotics*, 2025.

[2] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, "A soft robot that navigates its environment through growth," *Science Robotics*, vol. 2, no. 8, p. eaan3028, 2017.

[3] M. M. Coad, R. P. Thomasson, L. H. Blumenschein, N. S. Usevitch, E. W. Hawkes, and A. M. Okamura, "Retraction of soft growing robots without buckling," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2115–2122, 2020.

[4] J. Davy, T. P. Dean, N. J. Greenidge, B. Calmé, P. Lloyd, J. H. Chandler, and P. Valdastri, "Magnetic fluid-driven vine robots for minimally invasive tissue biopsy sampling," *Advanced Intelligent Systems*, vol. 7, no. 8, p. 2400827, 2025.

[5] N. G. Kim, N. J. Greenidge, J. Davy, S. Park, J. H. Chandler, J.-H. Ryu, and P. Valdastri, "External steering of vine robots via magnetic actuation," *Soft robotics*, vol. 12, no. 2, pp. 159–170, 2025.

A Low-Cost Smartphone-Based Diabetic Retinopathy Screening System Integrating AI Classification for Resource-Limited Settings

Ricardo J. Harrison[1], Gerardo Loza Galindo[2], Sasha-Gay Wright[1], Nikita J. Greenidge[2], Pietro Valdastri[2]

[1]*Faculty of Engineering, University of the West Indies, Mona*

[2]*University of Leeds, United Kingdom*
ricardocsn01@gmail.com

INTRODUCTION

Diabetic Retinopathy (DR) remains a leading cause of preventable blindness and a common microvascular complication of diabetes mellitus. In the Caribbean, particularly in Jamaica, the burden of diabetes and its complications remain high. Studies conducted at the University Hospital of the West Indies report a high prevalence of risk factors among patients with DR, including poor glycemic control, obesity, and elevated blood pressure [1]. Limited access to specialized ophthalmic equipment and trained personnel further exacerbates the challenge of early detection, especially in underserved and rural communities. Recent advances in artificial intelligence (AI), particularly deep learning, have enabled automated retinal image analysis, comparable to that of clinical experts [2]. Early convolutional neural network (CNN) models, such as Visual Geometry Group (VGG)-based architectures, achieved accuracies of 71%–83% [3], while more recent transfer learning approaches such as InceptionResNetV2 and DenseNet121 have reported accuracies up to 96.61% [4]. However, few studies integrate smartphone-based retinal imaging, embedded AI diagnosis, and physiological assessment into a unified screening system. To address this gap, this study presents an integrated screening system combining smartphone-based retinal imaging and AI-driven diagnosis using the EfficientNetV2-S architecture [5], offering an accessible, scalable, and cost-effective solution for early DR detection in resource-limited environments.

MATERIALS AND METHODS

The proposed screening framework integrates four principal components: (1) a public retinal image dataset, (2) an AI-based classification model, (3) a structured training and optimization strategy, and (4) a custom-developed low-cost hardware system for retinal imaging.

1. Retinal Image Dataset: The model was developed using the Diabetic Retinopathy 224×224 (2019) dataset [6], accessed via Kaggle. The dataset contains 3,662 color fundus images across five DR severity classes: No_DR, Mild, Moderate, Severe, and Proliferate_DR. It exhibits significant class imbalance, with the majority class containing 1,805 images and minority classes (Severe and Proliferate_DR) containing fewer than 300 images. A stratified 70/20/10 split was used for training, validation, and testing to ensure robust evaluation.

2. AI-Based Classification Model: At the core of the system is a transfer learning classifier based on the EfficientNetV2-S architecture [5], selected for its strong accuracy–computational efficiency trade-off. EfficientNetV2-S [5] incorporates progressive learning and compound scaling, enabling faster convergence and reduced computational cost compared to conventional CNNs. A custom classification head was added, consisting of global average pooling, batch normalization, fully connected layers with ReLU activation, dropout regularization, and a softmax layer for five-class prediction.

3. Training and Optimization: During training, a two-phase transfer learning strategy was used. The EfficientNetV2-S backbone was initially frozen to train the classification head, followed by gradual unfreezing of the upper 60% of layers for fine-tuning with a reduced learning rate. To address class imbalance, minority classes were oversampled, with sample weights applied via the tf.data pipeline for gradient scaling, and focal loss was used to improve sensitivity to underrepresented classes. Photometric and affine augmentations were applied during training, while Test-Time Augmentation (TTA) was used at inference to improve prediction stability and robustness. Post hoc threshold optimization was performed on referral classes using validation-set ROC analysis to improve sensitivity without materially affecting overall performance.

4. Low-Cost Hardware System: The trained classification model was embedded within a complete hardware–software pipeline comprising a custom-designed 3D-printed retinal imaging enclosure (Figure 2), an ESP32 and MAX30102-based physiological monitoring module (heart rate, SpO_2, temperature, and blood pressure), and a mobile/web application enabling real-time data acquisition and cloud-based inference. The retinal imaging enclosure was designed to align a 20D lens with a smartphone camera to enable fundus image acquisition. Proper optical alignment depends on the distance from the patient's eye (d_o) and the distance between the lens and the camera sensor (d_i), maintaining focal length (f) according to the thin-lens equation:

$$\frac{1}{f} = \frac{1}{do} + \frac{1}{di}$$

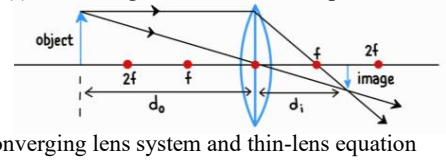


Figure. 1. Converging lens system and thin-lens equation

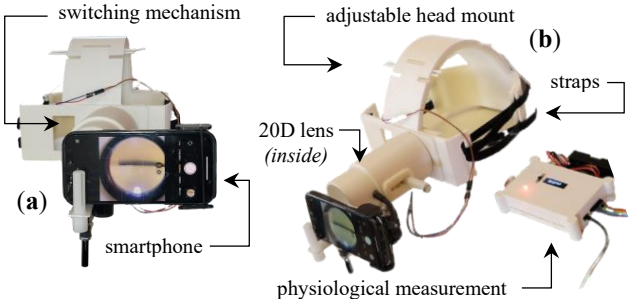


Figure. 2. Showing retinal imaging system from (a) front view and (b) top right view

RESULTS

The results are divided into two main components: (1) Model's Performance and (2) Prototype Performance.

1. Classification Model's Performance: After training, the model achieved a validation accuracy of 78.66% and a test accuracy of 78.86% on the test dataset. The application of TTA further improved test accuracy to 82.11%, with a corresponding weighted F1-score of 0.82, as summarized in Table 1. The confusion matrix in Figure 3 provides a detailed breakdown of classification performance across all five classes. The model performed well on the No_DR class, with most samples being correctly identified. Performance across the remaining classes varied, with some misclassification observed between adjacent severity levels, reflecting the visual similarity between consecutive stages of DR progression. To reduce the clinical risk of missing referral-stage disease, post hoc threshold optimization was performed using validation-set ROC analysis. A Youden-optimal threshold ($\tau = 0.261$) improved Moderate DR recall from 61.4% to 79.2% with minimal impact on overall accuracy (82.11% to 81.57%). Severe recall remained limited due to dataset constraints. At the binary referral level, the model achieved an AUROC of 0.978 and a sensitivity of 0.993, indicating reliable identification of referral cases regardless of severity grading.

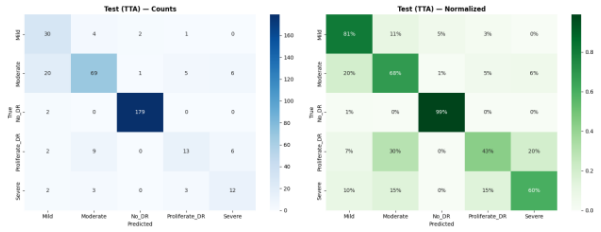


Figure. 3. Confusion matrix of model performance on the test dataset with TTA. Rows: true labels; columns: predicted labels.

Model Configuration	*Ac (%)*	*W_P*	*W_R*	*W_F1*
Validation	78.66	0.79	0.79	0.79
Test (No TTA)	78.86	0.80	0.79	0.79
Test (With TTA)	82.11	0.83	0.82	0.82
Test (TTA + Threshold)	81.57	0.81	0.82	0.80

Table. 1. Average metrics across model configurations: Accuracy (Ac), Weighted Precision (W_P), Weighted Recall (W_R), and Weighted F1-score (W_F1).

2. Retinal Imaging Prototype Performance: The retinal imaging enclosure enabled the capture of fundus images using a smartphone-based system. Figure 2 illustrates the design of the imaging device, while Figure 4 presents sample images captured during testing. The system was able to capture images at different stages of alignment, including initial positioning and fully illuminated retinal images. The optical configuration, incorporating a 20D lens positioned approximately 4–5 cm from the eye, allowed for image acquisition. Both simulated retinal images (from a model eye) and real retinal images were successfully captured using the prototype as shown in Figure 4.

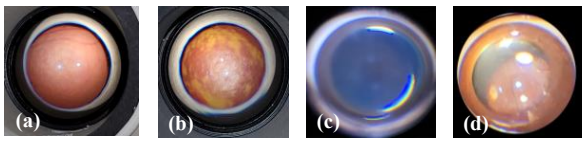


Figure. 4. Retinal images captured by the prototype: (a) healthy model eye, (b) unhealthy model eye, (c) lens alignment at the pupil, (d) real retinal image after illumination.

CONCLUSION AND DISCUSSION

This study presents an integrated screening system for DR tailored to resource-limited settings. By unifying smartphone-based retinal imaging, AI classification, and physiological monitoring within a single platform, it extends beyond conventional single-modality approaches. The EfficientNetV2-S model achieved 82.11% test accuracy with TTA and a weighted F1-score of 0.82. Threshold optimization improved referral-stage sensitivity with minimal impact on overall performance, supporting the system's potential for earlier intervention and reducing preventable vision loss.

This system addresses a critical gap in developing regions, where ophthalmic services and fundus imaging equipment are concentrated in urban centers, limiting access for high-risk populations. By integrating a low-cost 3D-printed enclosure with embedded AI analysis, the platform reduces reliance on specialized equipment and expert interpretation, while the physiological module supports broader cardiometabolic risk assessment. Future work will focus on local dataset development, optical refinement, and clinical validation to strengthen generalizability and real-world deployment readiness.

REFERENCES

[1] Foster T, Mowatt L, Mullings J. Knowledge, Beliefs and Practices of Patients with Diabetic Retinopathy at the University Hospital of the West Indies, Jamaica. Journal of community health. 2016;41(3):584–92.

[2] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA: the journal of the American Medical Association. 2016;316(22):2402–10.

[3] Nguyen QH, Muthuraman R, Singh L, Sen G, Tran AC, Nguyen BP, et al. Diabetic retinopathy detection using deep learning. In: Proceedings of the 4th International

Conference on Machine Learning and Soft Computing (ICMLSC 2020); 2020 Jan 17–19; Haiphong City, Vietnam. New York: ACM; 2020. p. 103–7.

[4] Guefrachi S, Echtioui A, Hamam H. Diabetic Retinopathy Detection Using Deep Learning Multistage Training Method. Arabian journal for science and engineering (2011). 2025;50(2):1079–96.

[5] Tan M, Le QV. EfficientNetV2: Smaller Models and Faster Training. 2021;

[6] Diabetic Retinopathy 224x224 (2019 Data) [Internet]. Kaggle; 2019 [cited 2026 Apr 17]. Available from: https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-2019-data

Magnetic Catheter Delivery with a Pressure-Lockable Hybrid Continuum Robot

Burak Ozdemir, Muhammad A. Khalid, Benjamin Calmé, Pietro Valdastri, and James H. Chandler

School of Electronic and Electrical Engineering, University of Leeds, UK

INTRODUCTION

Endoluminal procedures require flexible instruments capable of navigating confined and tortuous anatomical pathways while maintaining sufficient dexterity for precise positioning of the distal tool [1]. In procedures such as endoscopic retrograde cholangiopancreatography (ERCP), access to target regions can be challenging due to anatomical curvature, limited working space, and the need to orient tools toward small or difficult-to-reach ductal openings [2]. Continuum robots (CRs) offer an attractive solution because their compliant bodies can bend through constrained environments and provide shape adaptability during navigation [1]. However, conventional continuum deformation typically produces smooth bending along the robot's length, making it difficult to achieve highly localized or tight changes in direction near the target site. A robotic platform capable of switching between broad continuum bending and more localized delivery configurations may therefore provide a useful route toward adaptable endoluminal tool deployment [3].

In this work, we present a magnetic catheter delivery system comprising a hybrid tendon-pneumatic CR for coarse positioning and a magnetic soft robot (SR)/catheter for distal fine steering, as illustrated in Fig. 1A. Building on previously reported hybrid tendon-pneumatic CR principles [4], the system incorporates a central working channel for catheter insertion and deployment. Pneumatic pressurization provides both axial deployment of the hybrid CR and increased frictional contact between the compliant inner tube and the catheter, with the latter enabling reversible pressure-based catheter locking, while tendon actuation controls body bending. The system architecture, prototype, design parameters, and pressure-driven extension are shown in Fig. 1B-D. Unlike a fixed catheter guide, the hybrid CR can change the catheter exit position and orientation before the deployed magnetic SR is steered using an external magnetic field. This study demonstrates the preliminary feasibility of this coarse-to-fine approach through pressure-driven deployment, tendon-based body steering, magnetic distal actuation under different deployment conditions, and combined operation of both robotic components.

MATERIALS AND METHODS

The proposed hybrid CR consists of a pressure-actuated soft body with a central working channel, four tendon pathways, and a 0.6 mm outer-diameter magnetic SR/catheter delivered through the working channel.

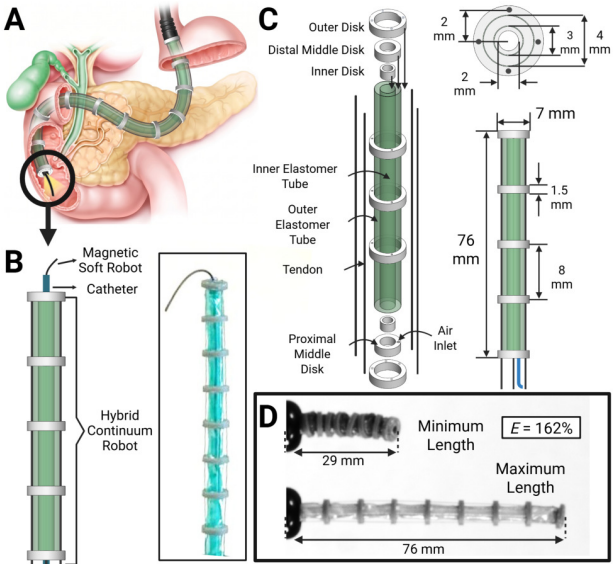


Fig. 1: Design concept and prototype of the proposed magnetic catheter delivery system. (A) ERCP-inspired application scenario, where the hybrid CR provides steerable access and delivers a magnetic SR/catheter toward the target region. (B) Overall system architecture showing the hybrid CR, catheter, magnetic SR, and fabricated prototype. (C) Exploded-view assembly, component layout, and main design parameters of the hybrid CR. (D) Experimental images showing the hybrid CR in its minimum and maximum length configurations, achieving 162% elongation from 29 mm to 76 mm. Created in BioRender. Ozdemir, B. (2026) https://BioRender.com/evakopc

The body was fabricated from 38 μm-thick TPE film (Stretchlon 200, Airtech Advanced Materials, UK) using heat-pressing and laser cutting to form outer and inner elastomer tubes with diameters of 4 mm and 2 mm, respectively, creating a pneumatic chamber around the working channel [5]. SLA-printed disk components were bonded to the elastomer tubes using cyanoacrylate glue, and four 0.2 mm NiTi tendons were routed through the disks' alignment holes and fixed distally to actuate whole-body bending across the eight-segment structure. A proximal silicone tube provided pneumatic access. The exploded-view assembly and key design parameters are shown in Fig. 1C.

Experiments were conducted to evaluate pressure-driven deployment, tendon-based robot bending, magnetic SR steering, and combined actuation. Magnetic actuation was performed using an electromagnetic micromanipulation system (MFG-100, MagnebotiX, Switzerland), with the magnetic field magnitude fixed at 20 mT while the field direction was varied. Tendons were actuated using minia-

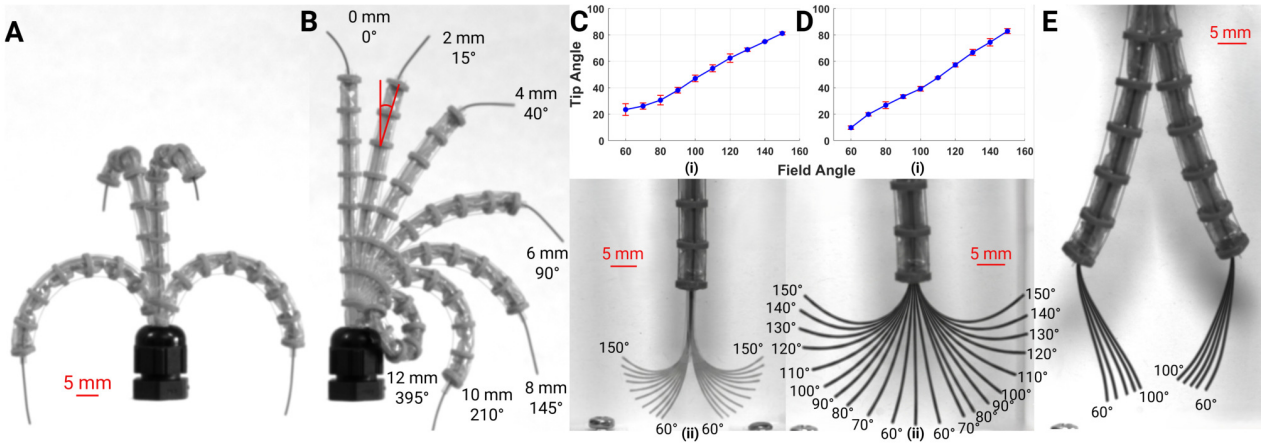


Fig. 2: Experimental characterization of the magnetic catheter delivery system. (A) Representative workspace of the hybrid CR under tendon-driven bending. (B) Tendon-actuated bending of the hybrid CR with measured tip angles for different tendon displacements. (C) Magnetic SR actuation after catheter deployment from the robot tip: (i) tip angle response with respect to the applied magnetic field angle, and (ii) corresponding experimental image sequence. (D) Magnetic SR actuation directly from the distal end of the hybrid CR: (i) tip angle response with respect to the applied magnetic field angle, and (ii) corresponding experimental image sequence. (E) Combined actuation of the hybrid CR and magnetic SR, demonstrating integrated body reconfiguration and distal magnetic steering.

ture linear actuators (L12-100-50-12-I, Actuonix Motion Devices) controlled through a cDAQ-based interface, while pneumatic actuation was regulated using pressure and vacuum regulators (SMC-ITV0010-3BL and SMC-IRV20-C08, respectively) and a 3/2 solenoid valve (RS PRO G-01V). A high-resolution camera (Basler aCA2040-120um) was positioned orthogonally to the robot, and recorded images were post-processed in ImageJ to extract backbone curves, tip positions, and tip angles.

RESULTS AND DISCUSSION

The fabricated hybrid CR demonstrated pressure-driven length modulation between its compact and deployed configurations. As shown in Fig. 1D, the hybrid CR extended from 29 mm to 76 mm, corresponding to a 162% elongation. This behavior supports compact insertion followed by pneumatic deployment, which is useful for endoluminal access where the available entry space is limited. Tendon-driven actuation enabled the robot body to generate multiple bending configurations and different delivery shapes. The representative workspace in Fig. 2A illustrates the segmented bending capability of the robot, while Fig. 2B shows the measured tip angle response for different tendon displacements. Increasing tendon displacement produced larger tip angles (up to 395°), confirming controllable body-level steering for coarse positioning of the catheter delivery point.

The magnetic SR was characterized under different deployment conditions. In the first configuration, the catheter was extended from the distal end of the hybrid CR, allowing the magnetic SR to be actuated after deployment. As shown in Fig. 2C, the magnetic SR followed the applied magnetic field direction, with the measured tip angle varying as a function of the field angle. In the second configuration, magnetic actuation was performed directly at the distal end of the hybrid CR, as shown in Fig. 2D. These results demonstrate distal fine steering from different delivery configurations, with magnetic SR tip angles of up to 85°. Finally, combined actuation of the hybrid CR and the magnetic SR was demonstrated, as shown in Fig. 2E. This highlights the main functionality of the proposed system: the hybrid CR provides body-level positioning, while the magnetic SR provides local distal steering from the selected delivery configuration. Compared with a fixed catheter guide, this approach allows the delivery path and distal actuation point to be adapted according to the target location and available anatomical workspace. Overall, these preliminary results demonstrate the feasibility of integrating pressure-driven deployment, tendon-driven body steering, and magnetic distal steering within a single catheter delivery platform. This coarse-to-fine architecture enables reconfigurable catheter positioning through the hybrid CR while providing localized distal steering using the magnetic SR. Future work will focus on modeling framework development, quantitative characterization of locking strength, workspace reconfiguration, targeting accuracy in ERCP-inspired phantoms, and closed-loop magnetic catheter control.

REFERENCES

[1] B. Ozdemir, M. A. Khalid, M. Chauhan, P. Valdastri, and J. H. Chandler, "Hybrid continuum robot designs and architectures for healthcare applications," *Advanced Robotics Research*, p. e202500177, 2026.

[2] B. M. Fung, T. C. Pitea, and J. H. Tabibian, "Difficult biliary cannulation in endoscopic retrograde cholangiopancreatography: an overview of advanced techniques," *European medical journal. Hepatology*, vol. 9, no. 1, p. 73, 2021.

[3] P. A. York, P. J. Swaney, H. B. Gilbert, and R. J. Webster, "A wrist for needle-sized surgical robots," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 1776–1781.

[4] B. Ozdemir, M. A. Khalid, P. Valdastri, and J. H. Chandler, "A hybrid tendon-pneumatic continuum robot with pressure-driven length modulation and tool-channel locking," in *2026 IEEE 9th International Conference on Soft Robotics (RoboSoft)*, 2026, pp. 950–955.

[5] J. Davy, T. P. Dean, N. J. Greenidge, B. Calmé, P. Lloyd, J. H. Chandler, and P. Valdastri, "Magnetic fluid-driven vine robots for minimally invasive tissue biopsy sampling," *Advanced Intelligent Systems*, vol. 7, no. 8, p. 2400827, 2025.

A Miniaturized Everting Robot with a Large Working Channel for Ureteroscopy Using Twisted Tail Pushing

Juliana Dreyer[1], Cédric Girerd[1], Cécile Manceau[2], Nicolas Abdo[2], and Nabil Zemiti[1]

[1]*LIRMM, University de Montpellier, CNRS, Montpellier, France.*
[2]*Department of Urology, Montpellier University Hospital, Montpellier, France.*

INTRODUCTION

Soft robots have gained significant attention for their inherent compliance and adaptability in unstructured environments. Among them, everting vine robots form a subgroup of soft, bio-inspired robots [1]. They grow from their tip via material eversion, where internally stored tail material turns inside out at the robot tip due to internal pressure, to form the robot body. This mechanism eliminates relative movement between the robot body and its environment, canceling both friction and shear forces.

Many applications require a working channel that provides access from the base to the robot tip for tool deployment. A common solution involves passing a tube through the tail to prevent it from collapsing due to the robot's internal pressure. However, the working channel tends to move with the tail due to friction. One existing strategy to mitigate this issue is to scrunch the tail material at the robot's distal extremity, reducing the length over which friction acts [2]. This solution becomes ineffective as the working channel diameter approaches the vine body diameter, reducing the radial storage space for the tail.

Ureteroscopy is a small-scale application where a ureteroscope is deployed through the ureter and retracted multiple times to break and extract kidney stones. To limit friction, current procedures involve the use of Ureteral Access Sheaths (UASs), which facilitate access to the kidney. Most UASs have inner and outer diameters of 10 Fr (3.3 mm) and 12 Fr (4.0 mm), respectively, and lengths of 350 mm for females and 450 mm for males. However, insertion of a UAS can damage the ureter due to large shear forces [3], leading to ureteral rupture in extreme cases.

In this work, we propose to replace the UAS by a novel vine robot design. The challenge lies in developing a robot with a large working channel to body diameter ratio and with significant growth length. We overcome friction between the working channel and the tail by pushing the scrunched and twisted tail material. This effectively decouples growth pressure from deployed length, removing a key limitation of existing designs. Finally, we fabricate a prototype and conduct experiments to validate the concept.

MATERIALS AND METHODS

The robot is composed of two tubes: the working channel and the tail pushing tube. The tail pushing tube pushes the tail towards the tip of the working channel, making the required pressure independent of the deployed length. The tail is twisted, limiting the outer storage diameter and preventing friction between the vine tail and vine body. Eversion of the twisted tail requires rotation of the tubes to unwind the tail and form a straight body. The process is illustrated in Fig. 1.

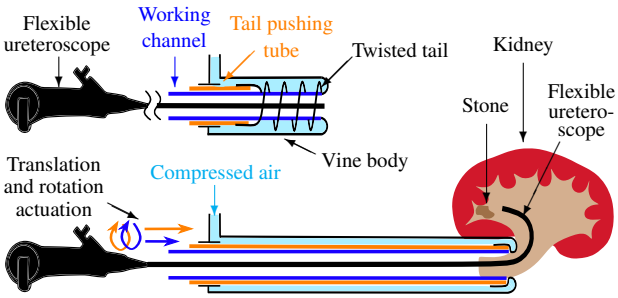


Fig. 1: Concept illustration of a vine robot as an ureteral access sheath. From the initial configuration (top), the vine robot is pressurized and the working channel and the tail pushing tube are both actuated in translation and rotation until the robot is fully deployed (bottom).

Fig. 2 shows the free body diagram of the forces acting on the robot. The system of equations governing the pressure, the pushing force on the tail pushing tube and working channel is given by Eq. (1):

$$\begin{cases} P_{grow} = \dfrac{F_y}{\pi(R^2 - r^2) - 4\mu_{VT/WC}\pi r z_s} \\ F_{TPT} = \mu_{VT/WC} P_{grow} 2\pi r z_t \\ \qquad + F_{VT/VB} + F_{TPT/VB} + F_{TPT/WC} \\ F_{WC} = -\mu_{VT/WC} P_{grow} 2\pi r (z_t + z_s) - F_{TPT/WC} \end{cases}, \quad (1)$$

where F_y is the eversion force, R is the robot inner radius and r is the working channel outer radius, $\mu_{VT/WC}$ is the friction coefficient between the vine tail and the working channel, z_s and z_t are the straight and the twisted storage lengths, $F_{VT/VB}$ is the friction of the vine body on its tail, $F_{TPT/VB}$ is the friction of the vine body on the tail pushing tube and $F_{TPT/WC}$ is the friction of the working channel on the tail pushing tube.

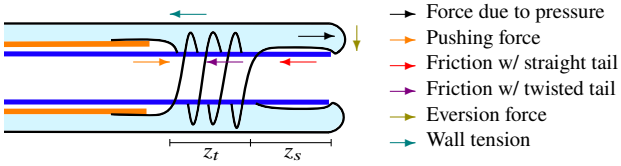


Fig. 2: Free body diagram of forces acting on the robot body and tail. The vine robot is illustrated with partially straight and partially twisted tail, allowing control over the pressure required to grow.

This work was supported by the ANR under Grant ANR-24-CE33-0538.

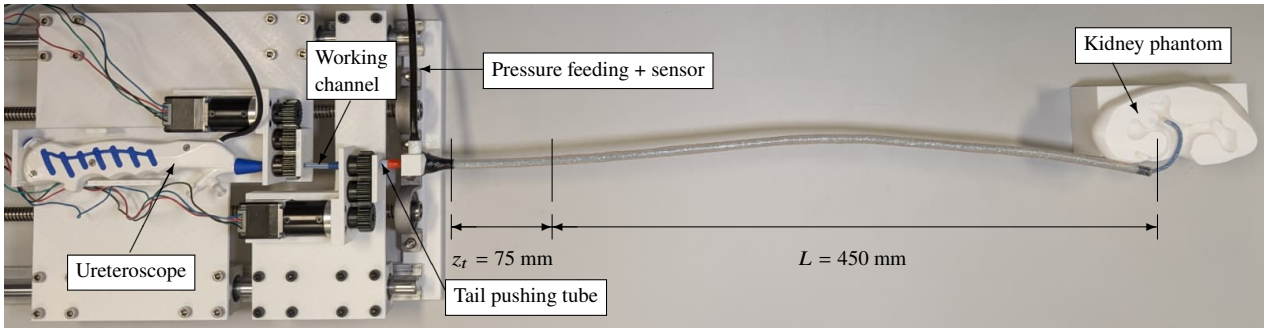


Fig. 3: Photo of the fabricated prototype. The working channel and tail pushing tube are translated and rotated by two individual platforms, located at their bases. The ureteroscope can be manipulated independently of the robot. In this experiment, the twisted material storage length is z_t = 75 mm, allowing the robot to grow L = 450 mm.

In this work, we focus on the case where $z_s = 0$, leading to $P_{grow} = \frac{F_y}{\pi(R^2-r^2)}$ in Eq. (1), and $F_{VT/VB}$, $F_{TPT/VB}$ and $F_{TPT/WC}$ are considered negligible. Control over the robot's length can be described by Eq. (2):

$$\Delta L = \Delta N \sqrt{(dz_t)^2 + (2\pi r_m)^2}, \tag{2}$$

where $\Delta N = \frac{\Delta\theta}{2\pi}$ is the number of tail turns, $dz_t = \frac{z_t}{N}$ is the lead of the tail and $r_m = \frac{R+r}{2}$ is the storage area mean radius. Assuming the tail has a constant lead, while the vine body and the working channel advance of ΔL, the advancing of the tail pushing tube is given by $\Delta L + \Delta z_t$, where Δz_t is the variation of twisted material storage length. The rotation and translation control parameters are described in Eq. (3):

$$\frac{\Delta z_t}{\Delta L} = \frac{2Rt}{\epsilon(R^2-r^2)}, \quad \frac{\Delta N}{\Delta L} = \frac{\sqrt{1-\left(\frac{\Delta z_t}{\Delta L}\right)^2}}{\pi\,(R+r)}, \tag{3}$$

where t is the vine material thickness and ϵ is the twisted tail compression ratio [2].

To test the concept, the robot body is fabricated from a 6.86 mm diameter and 50 µm thick PTFE Sub-Lite-Wall™ Liner (Zeus Company LLC). In order to find the eversion force, which depends only on the robot body material and its thickness [2], the pressure to grow this tube without a working channel and on a straight section was measured to be 32 kPa, leading to F_y = 1.18 N. The tail is then taped to a 4.72 mm inner and 5.72 mm outer diameter PTFE tail pushing tube. Another PTFE tube of 3.38 mm inner and 3.78 mm outer diameter, the working channel, is inserted inside the tail pushing tube. The ratio between the robot body diameter and the outer working channel diameter equals 0.55. When considering the working channel inner diameter, the ratio becomes 0.49. Finally, the robot body is everted and the tail twisted. Four motors control the translation and the rotation of the tubes. Fig. 3 shows the experimental setup.

During setup, the storage length was measured to find $\frac{\Delta z_t}{\Delta L}$ = 0.167. This gives a scrunching ratio of ϵ = 0.25 and a theoretical number of tail turns of $\frac{\Delta N}{\Delta L}$ = 0.059. In practice, $\frac{\Delta N}{\Delta L}$ = 0.044 was found. This difference can be explained by the approximate modeling of the mean storage radius, its variations in practice and the difficulty in measuring the number of turns while twisting the tail by hand.

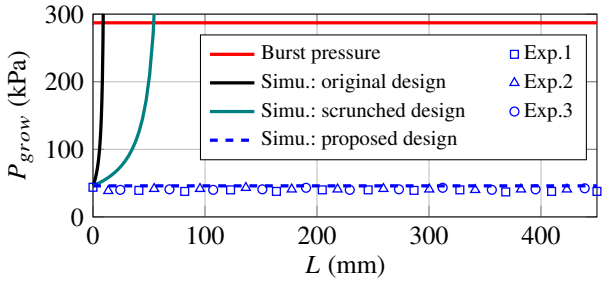


Fig. 4: Graph of deployed length vs pressure for different designs at prototype dimensions, showing that the proposed design removes length-dependent pressure increase.

Using ϵ = 0.25, F_y = 1.18 N, and $\mu_{VT/WC}$ = 0.1, the plot in Fig. 4 compares different designs and shows the experimental results of the proposed design. The original design (no material scrunching) and scrunched design from [2] reach the material burst pressure for deployed lengths of 9.1 and 54.4 mm, respectively. In contrast, our proposed concept of twisted-tail pushing deploys at a constant pressure of 46 kPa for a length of 450 mm, covering our application needs. The experimental results follow the trends predicted by the model, particularly regarding the independence of growth pressure from deployed length.

CONCLUSIONS AND DISCUSSION

This work demonstrates that twisting and actively pushing the tail enables miniaturized vine robots to deploy over long distances independently of internal pressure. This removes the coupling between deployment length and burst pressure that limits existing designs with large working channels, and opens the use to multiple applications including ureteroscopy. Future work will focus on further decreasing the robot diameter and on robot retraction, which remains a key challenge for clinical transfer.

REFERENCES

[1] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, "A soft robot that navigates its environment through growth," *Science Robotics*, vol. 2, no. 8, 2017.

[2] C. Girerd, A. Alvarez, E. W. Hawkes, and T. K. Morimoto, "Material scrunching enables working channels in miniaturized vine-inspired robots," *IEEE Transactions on Robotics*, vol. 40, pp. 2166–2180, 2024.

[3] S. N. Ali, A. McCormac, S. Saadat, S. A. M. Lavasani, A. D. Cumpanas, M. C. Hernandez, M.-C. Vu, A. S. Afyouni, M. O'Leary, J. Altamirano-Villarroel *et al.*, "Ureteral access sheath deployment: Understanding the force tolerance of the human ureter and the development of ureteral injury during endoscopic surgery," *Journal of Endourology*, p. 08927790251400350, 2025.

Towards Miniature Soft Vine Robots for Ultrasound Imaging of the Uterine Cervix

A. Yanez, N. J. Greenidge, B. Calme, J. Davy, P. Valdastri, J. H. Chandler, V. Pensabene, and J. Avery

School of Electronic and Electrical Engineering, University of Leeds

INTRODUCTION

Minimally invasive access to the uterus remains a persistent challenge in gynecology. Current endoluminal methods rely on rigid hysteroscopes or dilators that force passage through the cervix, a narrow, compliant, and sensitive structure. These probes are limited by discomfort, restricted maneuverability, and an inability to conform to the delicate anatomy. Consequently, they hinder the direct visualization required for early diagnosis of conditions such as endometriosis, adenomyosis, and cancer [1].

Soft growing robots, often referred to as "vine robots", navigate via tip extension driven by internal fluid pressure. Their unique kinematics allow them to traverse narrow lumens, minimizing shear forces along their length [2]. When used as medical catheters, they reduce tissue interaction forces, thereby enhancing safety during interventions. Nonetheless, their clinical translation depends on functionalizing these systems with imaging.

In this work, we present a millimeter-scale water-driven soft everting vine robot with integrated ultrasound imaging (Fig. 1). The design and fabrication strategies are introduced for attaching an Intravascular Ultrasound (IVUS) probe as the tail of a 5 mm vine robot, using a vascular imaging catheter for cervical imaging. Experimental results in a female reproductive system phantom demonstrate the successful navigation of the robot in the tight lumen while collecting ultrasound images. Our design is intended to facilitate navigation of delicate tissue for safe gynaecological diagnostics.

DESIGN AND FABRICATION

The vine robot was fabricated from two 38 µm thick Thermoplastic Polyurethane (TPU) films (Stretchlon 200, Airtech Advanced Materials, UK). A heat press (Combo Heat Press, Creworks, USA) was used to soft-seal the layers for laser cutting (Falcon 2 Pro 22W, Creality, China) the tubular shape of the robot body [3].

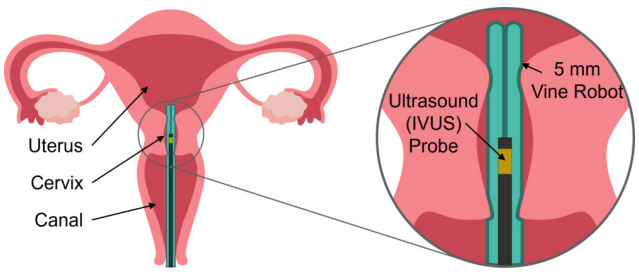


Fig. 1: The 5 mm vine robot carries an ultrasound (IVUS) probe through the cervix.

The phantom molds were generated in computer-aided design (CAD) from a GrabCAD geometry (Fig. 2 inset) scaled to anatomical dimensions, then cast in Dragon Skin^{TM} 10 (Smooth-On Inc, USA) silicone to simulate the geometric and mechanical properties of soft tissue. The tissue in the female reproductive system varies significantly depending on different stages, including age, gestational stage, and menstrual cycle [4]. Therefore, the material was selected to replicate the approximate resistance to access the uterus in the higher ranges of stiffness reported (Young's modulus $\sim 10^5$Pa) [4].

An IVUS probe (Visions PV.018 Digital IVUS Catheter, Philips Volcano, USA) was attached to the vine robot as the tail and clamped to a motorized linear rail to provide millimetric displacement control (Fig. 2), following our previous system [5]. The probe replaces the active channel and regulates eversion by relieving tension, thereby restricting independent motion with the vine. Unlike common pneumatic vine robots, this system uses pressurized water inside the robot lumen as the driving medium. The water serves as an essential acoustic-coupling transparent medium, minimizing acoustic impedance mismatch with soft tissue and preserving ultrasound image quality.

METHODS

Axial forces were measured over five insertions using the vine robot under tip eversion and a robotically inserted rigid probe ($Ø_{max}$ = 5 mm) to compare insertion mechanics.

Radial cross-sectional ultrasound slices were acquired (Volcano CORE Mobile System, Philips Volcano, USA),

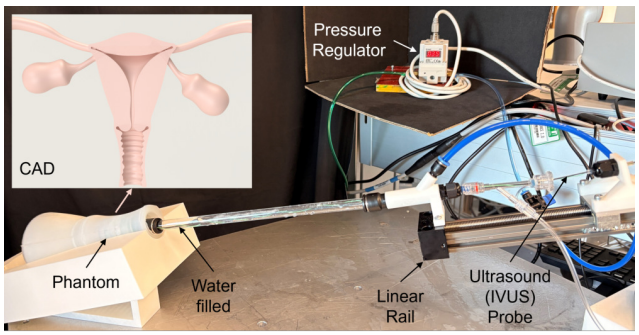


Fig. 2: Experimental setup for vine robot deployment with the IVUS probe into an anatomically scaled phantom of the female reproductive system (inset).

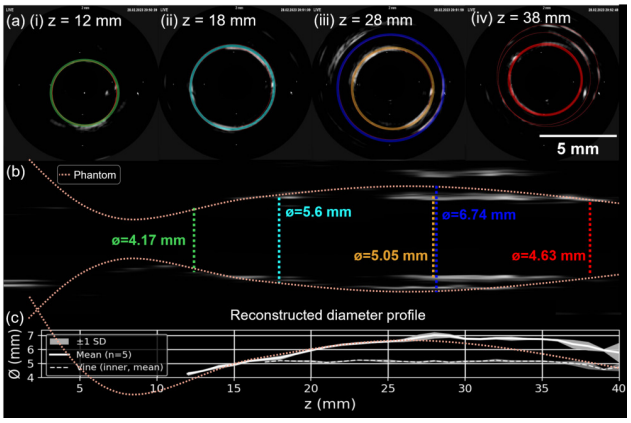


Fig. 3: Ultrasound-based reconstruction of the simulated cervical lumen. (a) Representative radial IVUS slices at insertion depths of (i) 12, (ii) 18, (iii) 28, and (iv) 38 mm. (b) Longitudinal multiplanar reconstruction, correlating the localized luminal diameters with the slices. (c) Mean diameter profiles of the vine robot and reconstructed cervical lumen, with shaded regions denoting ±1 SD (N=5). The cervical phantom profile is overlaid for reference.

while the probe was translated at 1 mm intervals along the longitudinal z-axis to a 40 mm insertion depth. The experiment was performed with the robot inserted into the water-filled phantom cervix and repeated five times. Noise was reduced by adjusting the settings for sensitivity, gain, and diameter to 3, 32, and 10 mm, respectively.

The slices were processed with custom Python scripts. First, each image was cropped to the region of interest with the dead zone ($r \sim 1$ mm) masked out. The localized diameters were manually identified in each slice and measured computationally, excluding partially visualized cross-sections. For one representative repetition, images were aligned to the centers of the smallest diameters. Segmented contours were finally stacked at 1 mm voxel resolution to reconstruct the lumen.

RESULTS AND DISCUSSION

The vine robot exhibited a maximum insertion force of 0.145 ± 0.012 N, compared with 3.221 ± 0.172 N for the rigid device, highlighting the reduced interaction forces enabled by tip eversion during cervical navigation.

Fig. 3 illustrates ultrasound-based shape reconstruction of the cervical path within the phantom model. In Fig. 3(a), cross-sectional scans show the measured circles at representative insertion depths across the z-axis. The first region, from $z = 1$ to (i) 12 mm, did not provide sufficient image contrast for reliable circular feature detection. The inner circles from (ii) $z = 18$ mm through $z = 36$ mm, including (iii) $z = 28$ mm, correspond to the vine robot diameter ($\varnothing = 5.09 \pm 0.05$ mm). This diameter decreases to $\varnothing = 4.56 \pm 0.06$ mm near the end, $z = 39$ mm.

Fig. 3(b) shows the sliced longitudinal y–z reconstruction obtained after rotating the image stack 60°clockwise around the z-axis, which provided the most complete cross-sectional representation along the z-axis. Most of the detected luminal edges appeared as fragmented contours due to signal dropout at certain angles. Therefore, Fig. 3(c) plots the two innermost extracted diameters, enabling a direct comparison between the smallest detected circle, denoting the vine, and the phantom profile. The largest diameter is observed around $z = 27$ mm, corresponding to the widest section of the cervical lumen.

Although the measured data capture the overall geometric trend of the lumen, a slight deviation remains (0.65 ± 0.06 mm, N=5), with the ground-truth profile positioned between the detected circles. This offset stems from using an IVUS catheter designed for vascular imaging, which can lead to geometric distortions in non-vascular environments. Furthermore, acoustic attenuation and boundary discontinuities caused luminal edges to appear as partial arcs, necessitating circular fitting to estimate the complete diameter.

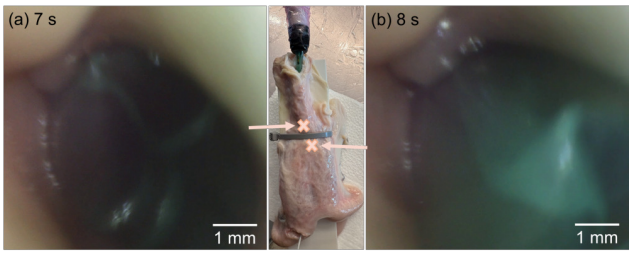


Fig. 4: Endoscopic visualization of robotic deployment during preliminary ex vivo tests on ovine cervix and uterus tissue. (a) The tip emerges from the cervix at 7 s. (b) At 8 s, the vine advances ~ 3 mm into the uterus.

Further preliminary ex vivo tests conducted in ovine cervix confirmed that the vine robot could deploy through cervical tissue and advance into the uterine cavity (Fig. 4). Corresponding ultrasound imaging, however, was limited by high tissue attenuation and reduced contrast, preventing reliable feature identification. These findings highlight the need for improved acoustic sensitivity and coupling for real-time imaging in biological tissue.

CONCLUSION

Overall, the proposed system demonstrates the feasibility of ultrasound-based cervical navigation using a soft vine robot while substantially reducing insertion forces compared with a rigid probe, providing a foundation for minimally invasive intrauterine interventions.

REFERENCES

[1] E. Pascoal *et al.*, "Strengths and limitations of diagnostic tools for endometriosis and relevance in diagnostic test accuracy research," *Ultrasound in Obstetrics and Gynecology*, vol. 60, pp. 309–327, 9 2022.

[2] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, "A soft robot that navigates its environment through growth," *Science Robotics*, vol. 2, p. 3028, 7 2017.

[3] J. Davy, T. P. Dean, N. J. Greenidge, B. Calmé, P. Lloyd, J. H. Chandler, and P. Valdastri, "Magnetic fluid-driven vine robots for minimally invasive tissue biopsy sampling," *Advanced Intelligent Systems*, p. 2400827, 2025.

[4] A. Baah-Dwomoh, J. McGuire, T. Tan, and R. D. Vita, "Mechanical properties of female reproductive organs and supporting connective tissues: A review of the current state of knowledge," *Applied Mechanics Reviews*, vol. 68, 11 2016.

[5] A. Yanez, J. Davy, J. H. Chandler, J. Avery, and P. Valdastri, "Miniaturized magnetic tip design for endoluminal vine robot navigation," *Advanced Robotics Research*, p. e202500188, 2 2026.

Improved Passive Navigation in Soft Growing Eversion Robots Through Embodied Intelligence via Constrictive Bands

Cem Suulker, Muhie Al Haimus, Thomas Mack, Mohammad Sheikh Sofla, Neri Niccolò Dei, Reza Kashef, S.M.Hadi Sadati, and Kaspar Althoefer

Centre for Advanced Robotics @ Queen Mary, School of Engineering and Materials Science,
Queen Mary University of London, United Kingdom.

INTRODUCTION

Eversion robots are soft growing robots that extend from their tip through fluidic pressure-driven eversion. Typically fabricated from thin flexible materials such as plastic films or coated fabrics, they can reach confined environments while exerting minimal interaction forces on their surroundings. These properties make them attractive for medical applications such as colonoscopy [1].

In constrained environments, eversion robots naturally follow paths of least resistance but struggle to negotiate sharp bends. Previous studies have shown that conventional designs fail below certain bending radii [2], whereas colonoscopy requires reliable traversal through multiple tight curvatures [1].

Existing navigation strategies can be broadly divided into active and passive approaches. Active steering approaches [3], [4] improve maneuverability but increase complexity and stiffness, compromising key features such as compliance and squeezability. Passive approaches are generally simpler but often rely on predefined shapes that are unsuitable for unknown environments [5].

This paper presents a passive steering mechanism for eversion robots, in which integrated constrictive bands across the length of the growing robot create localized hinge-like deformation zones with reduced bending stiffness, enabling sharper bending through environmental interaction (Figure 1).

MATERIALS AND METHODS

Fabrication: For this study, 0.05 mm TPU sheets were used as the base material for the eversion robot body. Constrictive TPU bands were laser cut and ultrasonically welded onto the sheet before cylindrical sealing and inversion. All prototypes were fabricated with a diameter of 40 mm. A 3D-printed alignment mask was used to improve fabrication repeatability.

Experiments: This study investigates how constrictive bands influence the bending stiffness of eversion robots. Multiple independently fabricated prototypes were tested while systematically varying band number, placement, width, and constriction ratio.

Mechanical characterization was performed using an Instron universal testing machine (Figure 2a), which applied lateral displacement to the robot tip while measuring reaction forces. Bending stiffness was quantified using force–displacement gradients.

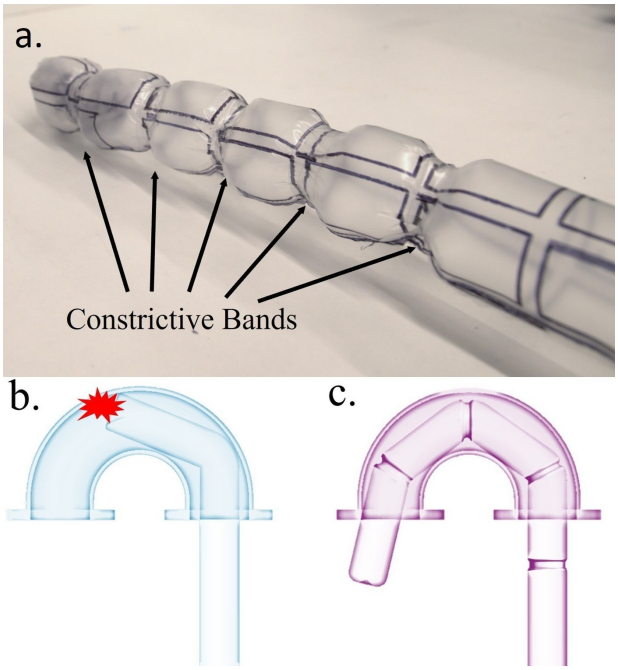


Fig. 1: a) Detailed view of the integrated constrictive bands. b) A standard eversion robot becoming stuck at a sharp bend. c) A band-integrated eversion robot successfully traversing the same bend through localized hinge-like deformation.

Case Study: To evaluate navigation capability in anatomically relevant conditions, standard and banded eversion robots were tested in silicone colon and abdominal phantoms replicating anatomical morphology (Figures 2c–h). The banded robot used multiple 10% constricting bands spaced 30 mm apart, selected based on characterization results. Internal pressure was controlled up to 8 kPa while robot motion and pressure were recorded.

RESULTS AND DISCUSSION

Across all experiments, constrictive bands consistently reduced bending stiffness and improved passive navigation capability. Increasing the number of bands produced a cumulative stiffness reduction (Figure 2b), although the effect gradually saturated at higher band counts. This suggests that adding bands progressively increases the number of compliant deformation zones, but that additional bands provide diminishing returns. Bands positioned further from

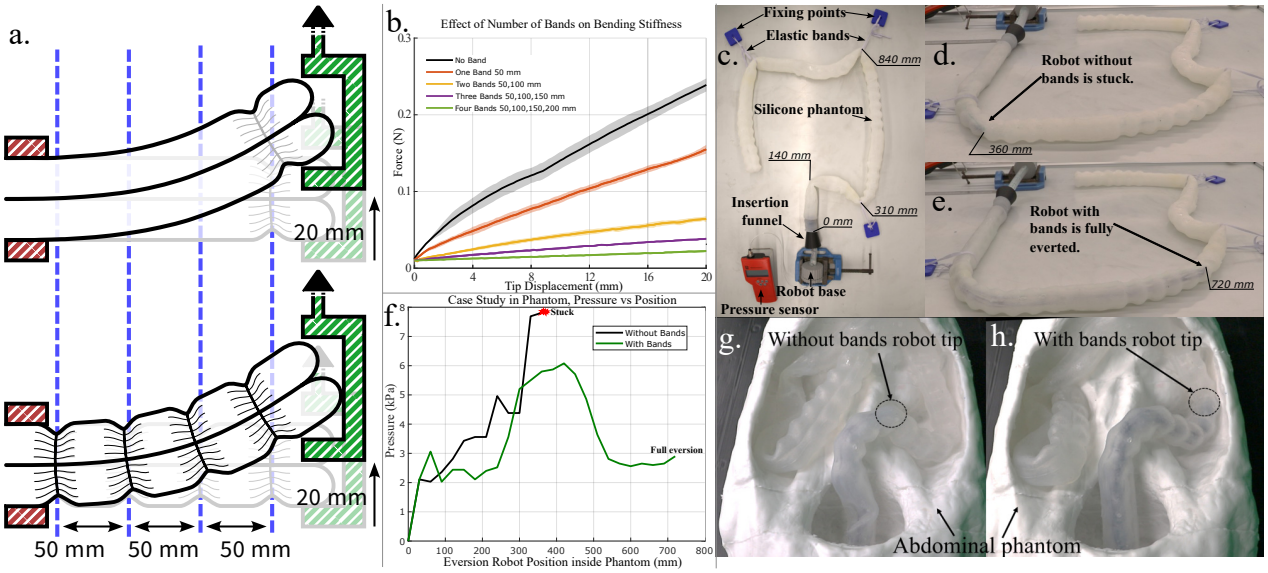


Fig. 2: Experimental setup for bending stiffness measurement (a) and corresponding results (b). Silicone phantom setup (c). Standard eversion robot stalls at a 90° bend (d), while the band-equipped robot successfully traverses it (e). Pressure–position results are shown in (f). In the abdominal phantom, the standard robot stalls at the first bend (g), whereas the band-equipped robot successfully navigates it (h).

the tip reduced bending stiffness more effectively, while increasing band width provided only a moderate benefit. These results indicate that band number and position are more influential design parameters than band width.

Experimental results demonstrate that constrictive bands reduce the robot's bending stiffness at the tip by up to 92%, enabling consistent traversal of 180° bends with a bending radius as low as 25 mm—notably lower than the 35 mm achievable by standard eversion robots under identical conditions. The substantial stiffness reduction demonstrates that the bands can introduce a hinge-like deformation behaviour without requiring additional actuators or rigid steering components.

Increasing constriction ratios enabled further reductions in bending stiffness but substantially increased the pressure required for eversion, producing a trade-off between compliance and actuation feasibility. Constriction ratios between 10% and 30% therefore provided the most practical balance between maneuverability and reliable eversion. This trade-off highlights that maximum compliance is not necessarily the optimal design objective, as excessive constriction can compromise the robot's ability to grow.

Comparisons with uniformly reduced-diameter robots showed that discrete banded structures can achieve lower bending stiffness than continuously reduced geometries. This indicates that localized wrinkling and geometric discontinuities contribute significantly to compliance enhancement, beyond the effect of diameter reduction alone. Thus, constrictive bands provide a means of retaining a relatively large robot diameter while locally achieving very low bending stiffness.

In both phantom setups, the standard robot progressed through straight sections but stalled at the first major sigmoid bend at pressures approaching 8 kPa (Figures 2d,g). In contrast, the banded robot successfully traversed the 90° bend and navigated additional downstream curvature at lower peak pressures (Figures 2e,h). These results demonstrate that the reduction in bending stiffness measured in the characterization experiments translates directly into improved passive navigation in compliant, colon-like environments, without requiring active steering mechanisms.

CONCLUSIONS

This work introduced constrictive bands as a simple structural modification for eversion robots that enables tunable reduction of bending stiffness and improved passive navigation in confined environments. By locally constraining the robot diameter, the bands induce hinge-like deformation without added actuation complexity or requiring much greater pressure. Experimental results demonstrated substantial stiffness reduction and improved traversal of sharp bends in colon-like phantoms, although full autonomous colon traversal remains an open challenge.

ACKNOWLEDGMENT

This work is supported by ERC grant EndoTheranostics, 101118626. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] C. Suulker *et al.*, "State of the art in soft eversion robots for colonoscopy: a review," *Progress in Biomedical Engineering*, 2026. [Online]. Available: http://iopscience.iop.org/article/10.1088/2516-1091/ae37b2

[2] D. A. Haggerty *et al.*, "Characterizing environmental interactions for soft growing robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 3335–3342.

[3] M. M. Coad *et al.*, "Vine robots," *IEEE Robotics & Automation Magazine*, vol. 27, no. 3, pp. 120–132, 2019.

[4] T. Mack *et al.*, "Efficient manufacturing of steerable eversion robots with integrated pneumatic artificial muscles," *Journal of Manufacturing and Materials Processing*, vol. 9, no. 7, p. 223, 2025.

[5] J. D. Greer *et al.*, "Robust navigation of a soft growing robot by exploiting contact with the environment," *The International Journal of Robotics Research*, vol. 39, no. 14, pp. 1724–1738, 2020.

Origami Polyimide Sleeves for Protection of Variable Length Soft Endoscopic Actuators

Thomas P. Dean[1,2], Benjamin Calmé[1], Burak Özdemir[1], Bao N. Nguyen[2], Pietro Valdastri[1], and James H. Chandler[1]

[1]*STORM Lab Leeds, University of Leeds*
[2]*School of Chemistry, University of Leeds*

INTRODUCTION

Due to the prevalence of endoscopic procedures, the transmission of infections and pathogens through patient interaction with endoscopic devices is a growing possibility [1]. Endoscopic devices often require low temperature sterilisation methods due to delicate components [2]. However, these methods are ineffective in comparison to heat-treatment which both penetrates further and is more effective at comprehensive disinfection [3]. The development of silicone based soft robotic endoscopes have been previously investigated with the goals of an improvement in patient comfort and rapid manufacturing. However, the material properties of silicone are impacted with an increase in temperature [4]. Therefore there is a need for a highly temperature stable solutions that can deliver the manoeuvrability of soft robotic devices and the resilience of traditional endoscopes. This work presents a heat stable protective sleeve for integration with soft robotic endoscopes, allowing both device protection and improved dexterity due to its structural compressibility and high torsional stiffness (Fig. 1).

MATERIALS AND METHODS

A. Sleeve Design and Integration

Polyimide (PI, Kapton® Goodfellow; thickness: 0.05mm) was chosen for the main material of a thermally stable protective sleeve due its resistance to temperatures up to 800°C, thereby indicating the resilience to the thermal conditions in an autoclave (<150°C). Origami cylinders were manually folded to a pattern designed in Fusion360. An Accordion fold pattern was chosen due to its high compressibility and high torsional stiffness. Unlike the popular Kresling folding, occurring through twist buckling of cylinders, the Accordion design resists axial torsion to promote stable bending during soft robot actuation. The design and parameters are presented in Fig. 2 and Table I. The pattern was transferred to a polyimide sheet using a ballpoint pen to indent the polyimide and define crease lines for easier folding. Folded sheets were rolled into cylinders and clamped to a tube. Temperature resistant adhesive (JBWeld HighHeat, RS Components) was subsequently applied along the overlapping seam.

The sleeve was integrated with a soft continuum device constructed from a multi-lumen silicone tube and containing three nitinol wires (0.4 mm) looped through and

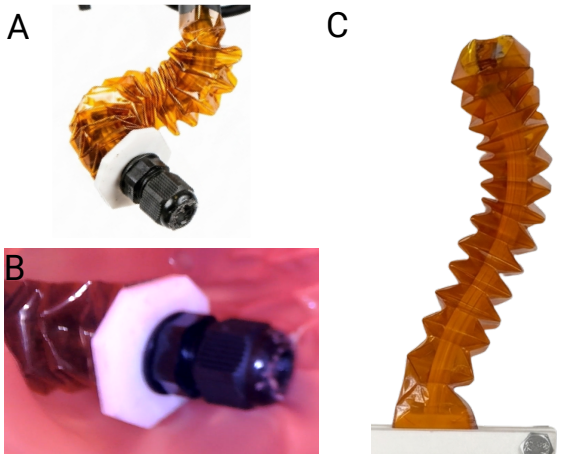


Fig. 1: The origami polyimide sleeve integrated with a tendon-driven soft continuum device, actuated in: A) Free-space B) A mock clinical simulator C) Free-space.

connected to 3D printed handles which allowed actuation over a three-dimensional workspace. To form a non-permanent attachment of the silicone tube to the polyimide sleeve, a PG7 cable gland was glued to one end of the polyimide sleeve to clamp the silicone tube. The other end of the sleeve was attached to another 3D printed part through which the silicone tube was threaded. Although a tendon actuated device was used here, alternative soft continuum robot actuation and/or design approaches (e.g. tip mounted magnet actuation with a tendon for axial constraint) could easily be integrated with the origami sleeve.

B. Mechanical Characterisation

Independent mechanical characterisation of the sleeve was conducted with a load cell (900 series S Beam, Load-CellShop) mounted on a linear stage (THORLABS Linear Translation Stage 150 mm). The stage was programmed to compress the sleeve to a length of 30 mm and axial force measurements were taken as the stage was set to travel 150 mm (extending the sleeve length to 180 mm). Demonstration of thermal stability of the PI sleeve was conducted via a thermal cycling approach whereby the sleeve was placed into an oven set to 200°C for one hour, before being left to cool and then three repeats of the mechanical characterisation experiment was conducted again for 3 cycles. Overall resulting in a maximum force during compression of 2.4 N, under tension of 2.6 N and across all repeats an s.d. of 0.09 N.

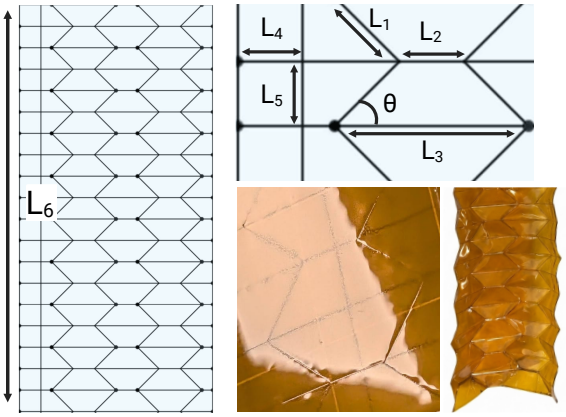


Fig. 2: Origami polyimide sleeve design and construction.

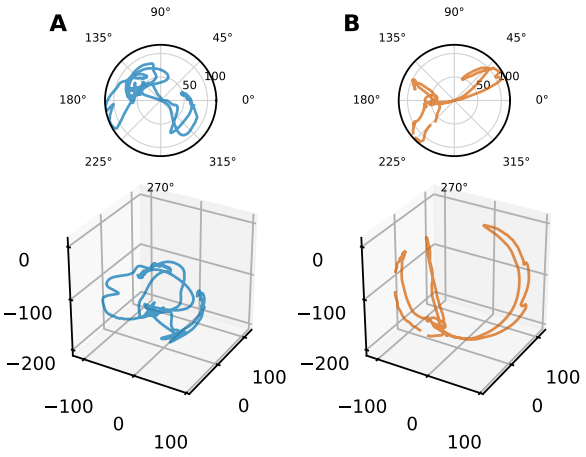


Fig. 3: A uniaxial trajectory for a 120 mm construction showing the least and most controllable assemblies of: A) the silicone tube alone and B) the silicone tube with a polyimide sleeve.

TABLE I: Accordion pattern origami design parameters.

Symbol	Value
L_1	14.14 mm
L_2, L_4, L_5	10 mm
L_3	30 mm
L_6	280 mm
θ	45°

C. Workspace and Navigation

Tip trajectories were tracked with a four camera Optitrack system and Motive 2.3 software to submillimeter accuracy (mean error: 0.0103mm). The device was deployed at three different lengths, 60 mm, 90 mm and 120 mm, and was controlled through uniaxial and multiaxial trajectories (Fig. 3). To demonstrate the torsional stiffening via the sleeve, a comparison was made between a stiffening back-bone and the PI sleeve. Therefore, four constructions of device were characterised, the multi-lumen silicone tube alone, the silicone tube with the sleeve, and both of these with a nitinol wire (0.6 mm) inserted through a central lumen (Fig. 1 A).

Arbitrary path navigation was demonstrated past two 3D-printed obstacles. The route comprised of an S-shaped path requiring the user to start from a fully compressed polyimide sleeve before deploying and actuating to navigate through a 120 mm route towards a target region. The sleeve-integrated device also demonstrate the possibility of retaining a continuous working channel. This was demonstrated through targeted fluid delivery, where coloured water was injected from the proximal end of the silicone tube and delivered to the distal end located in the target region. The channel was also used for the addition of stiffening backbones and connecting wires for tip based components.

XY tip-tracking was performed and an assessment was made between an optimal tip path and the actual path of the device. Disparities between the calculated and measured path were attributed to the manual actuation of the silicone tube and the duty cycling approach needed due to the opposing input forces of the lateral control and the extension of the device. It was also observed that the silicone tube buckled during extension past 120 mm (the relaxed length of the PI sleeve). Finally, a mock clinical demonstration was conducted in a stomach phantom. The device was inserted into the stomach and manual control via individual tendon actuation was demonstrated before the deployment of liquid through an internal working channel in the multi-lumen tube to simulate a targeted drug delivery (Fig. 1 B).

CONCLUSIONS AND DISCUSSION

This work presents a novel use of polyimide as an origami sleeve for use as a resilient outer layer for soft robotic endoscope devices. Through a careful choice of folding pattern, a highly compressible origami sleeve was achieved, resulting in a compression from a relaxed length of 120 mm to 30 mm with 2.4 N of axial force. Extension of the sleeve from relaxed length was conducted to 180 mm (60 mm displacement from relaxed state) and resulted in a tensile force of 2.6 N. These values showed negligible change after three repetitions of thermal cycling at 200°C for one hour. During the workspace characterisation, internal friction in the multi-lumen silicone actuator was minimised. This demonstrated that the PI sleeve improved controllability to a greater extent than a nitinol stiffening backbone. This work successfully constructed a compressible origami sleeve that aims to provide a sterilisable outer layer for surgical applications and to minimise the thermal degradation in soft robotic medical devices.

REFERENCES

[1] D. J. Weber, "Managing and preventing exposure events from inappropriately reprocessed endoscopes," *Infection Control & Hospital Epidemiology*, vol. 33, no. 7, pp. 657–660, 2012.

[2] S. Mattke, "Sterilization of endoscopic instruments," *JAMA*, vol. 313, no. 5, pp. 524–524, 2015.

[3] A. Pereira, G. Nakka, S. Gupta *et al.*, "Beyond ethylene oxide (eto): a comprehensive review of sustainable sterilization technologies for medical devices," *Biomedical and Therapeutics Letters*, vol. 12, no. 2, pp. 1161–1161, 2025.

[4] J. Lavazza, "A study on the mechanical behaviour of ecoflex 00-50 silicone elastomer," 2020.

Towards Automated Fabrication of Soft Robotic Endoscope Tips for Affordable Gastroscopy

Alexander Brown[1], Pietro Valdastri[1], and James H. Chandler[1]

[1]*School of Electronic and Electrical Engineering, University of Leeds*

INTRODUCTION

Upper gastrointestinal endoscopy procedures exceed 6.9 million annually in the USA alone, yet current flexible endoscopes present substantial challenges: reprocessing costs up to $324 per procedure, breakage rates of 1-10%, and cross-contamination risks [1]. Multi-chamber fluidic Soft Continuum Robotic (SCR) alternatives could address these issues, and have been extensively researched as a technology for flexible endoscopy due to their compliant tissue interaction [2]. However, many SCR approaches rely on complex manufacturing, which potentially limits their viability as scalable, cost-effective alternatives to traditional flexible endoscopes.

Traditional SCR fabrication involves casting liquid silicone into bespoke molds followed by manual demolding. Manufacturing multi-chamber SCRs with the complex internal geometry needed to induce desired motion often requires numerous fabrication process steps [3]. When applied at small scales, the procedural complexity intensifies and can act to limit actuator reliability and repeatability.

This work introduces an approach to automating the manufacturing process for monolithic soft robotic endoscope tips using synchronized helical core extraction (Fig. 1a,b). We modify the Parallel Helix Actuator (PHA) design [4] to incorporate a miniature camera and a light source (Fig. 1c), and develop a machine and tooling for automated core removal. The resulting system produces three-chamber actuators (13 mm diameter, 50 mm length) with >180° bending capability, establishing a pathway toward scalable SCR endoscope fabrication. We evaluate bending performance of the actuator and perform visual inspection within a stomach phantom.

MATERIALS AND METHODS

System Architecture

The fabrication system comprises four components enabling partially automated production of multi-chamber SCR actuators. First, a PHA design was modified to incorporate a 3 mm central channel for camera integration while maintaining three pneumatic chambers arranged 120° around the centerline (1 mm wall thickness, 13 mm overall diameter, and 50 mm length). Second, reusable Titanium helical cores (Grade 5 Ti6-Al4-V) were manufactured via direct metal laser sintering to provide the internal chamber geometry when overmolded during casting. The cores have parameters of 1.5 mm shaft diameter, 4 mm pitch, and a 40 mm helical length. Superelastic Nitinol wire extensions (0.75 mm diameter) were bonded to each core allowing for attachment to the extraction machine while transmitting the demolding forces.

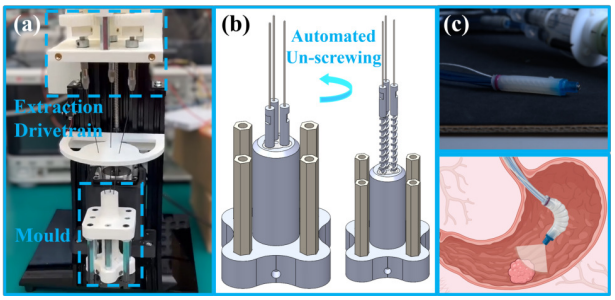


Fig. 1: Automated helical core extraction system and resulting SCR endoscopic tip. (a) Automatic core extraction machine with gearbox-driven synchronized rotation and vertical translation. (b) Before and after automated extraction of the helical cores. (c) Fabricated endoscope tip with integrated 3.9 mm diameter camera and light source (top), depiction of the SCR endoscope in upper gastrointestinal tract for diagnostic visualization (bottom).

Third, the three-part mold was developed, consisting of: (i) Base with 3 mm stem for forming the camera cable channel and features to encapsulate the interlocking thread for the camera housing unit; (ii) Casing defining external geometry and silicone injection port; (iii) End cap securing helical cores during casting with gas bleed holes. Fourth, the automated extraction machine provides vertical translation via a NEMA 23 stepper motor and Tr8x2 lead screw, with synchronized rotation via a NEMA 17 stepper motor driving a custom 3:1 gearbox. Three pin vises (162A; Starrett, USA) secure the cores, with microcontroller co-ordination ensuring 4 mm vertical translation per core rotation to match the helical profile.

Fabrication Protocol

After mold assembly and core alignment, 20 ml silicone elastomer (Ecoflex 00-30; Smooth-On, USA) is injected and cured (30 min, 60°C). The end cap is replaced with a locking plate securing the mold to the machine base. Nitinol wire extensions pass through alignment holes to the pin vises, and automated extraction executes 40 mm vertical translation with 10 simultaneous rotations, matching the pitch of the helical cores. Post extraction, manual mold disassembly retrieves the endoscope tip. A 3.9 mm diameter camera with integrated LED's (3.9mm 720P Endoscope Camera Module; DoTheCamera, China) inserts through the central channel, and interlocks with the housing unit. The full process can be seen in Fig. 2.

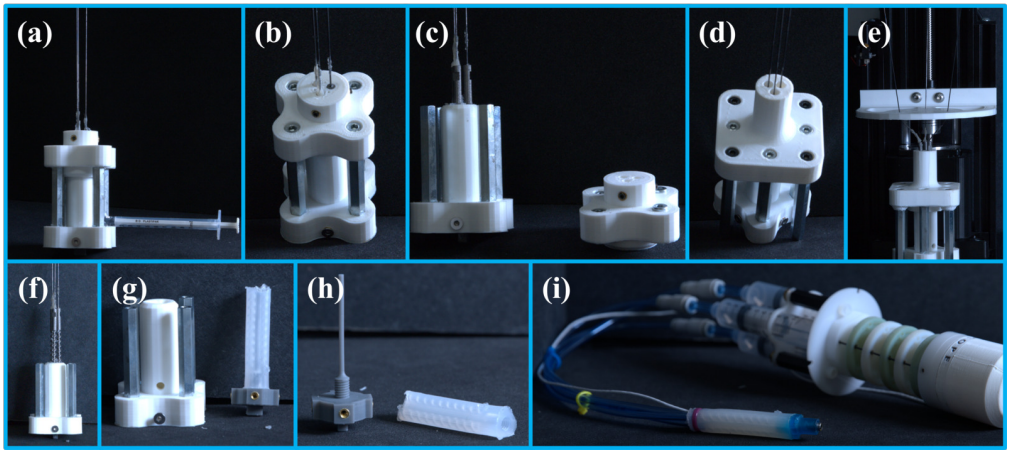


Fig. 2: Process for manufacturing. (a) SCR cast into three part mold. (b) SCR cured, showing alignment of helical cores. (c) Removal of end cap. (d) Locking plate secured to mold. (e) Automated extraction of helical cores from mold. (f) Removal of locking plate, showing vertical translation of helical cores post extraction. (g) Removal of base from the mold. (h) Removal of SCR from mold base. (i) Finished product with integrated camera and actuation unit.

Performance Evaluation

Workspace characterization was performed using electromagnetic tracking (Aurora, Northern Digital Inc, Canada), mapping the three-dimensional tip position during independent and combined chamber actuation (20-40 kPa). Each chamber was independently actuated by a solenoid valve (ITV0010, SMC Pneumatics), controlled via a data acquisition card (cDAQ, National Instruments, USA) following sinusoidal pressure profiles with 120° phase offsets to produce circular motion of the PHA (Fig. 3a). At 40 kPa, >180° bending was achieved where the pressure was limited to remain below the burst threshold. Phantom validation was performed in a silicone stomach phantom (Fig. 3b), simulating clinical Esophagogastroduodenoscopy procedures. A handheld syringe manipulation tool, based on the design in [5], was used to manually bend the tip during inspection based on a video feed from the integrated camera. In 2 minutes of actuation, no leaks or damage to the PHA was detected with a full range of motion demonstrated and key anatomical sites visualized.

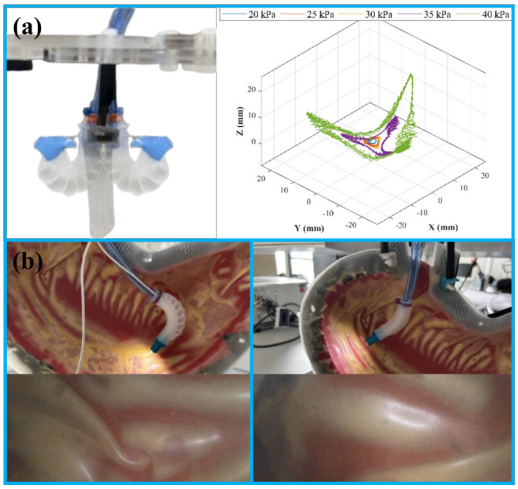


Fig. 3: Validation of the fabricated endoscope tip. (a) Workspace characterization setup (left) and three-dimensional tip trajectories at 20-40 kPa. (b) Phantom validation showing external maneuverability within the stomach phantom (top) and camera visualization (bottom).

CONCLUSIONS AND DISCUSSION

This work addresses the challenge of demolding soft robotic, small-scale, multi-chamber actuators by adapting a PHA design for automated extraction of the helical cores through synchronized rotation and translation using a custom machine. The resultant endoscope tip was evaluated through tip workspace characterization and stomach phantom validation. The automated process has the potential to reduce operator time and actuator variability compared to manual demolding.

Current limitations include the 13 mm diameter exceeding typical gastroscope dimensions (10 mm) and the manual steps in mold assembly and disassembly. Future work will address camera miniaturization for diameter reduction, full process automation including silicone injection, and extraction repeatability through the assessment of multiple PHA core extractions. This work establishes a pathway toward scalable manufacturing of disposable soft robotic endoscopic tools, addressing reprocessing costs and cross-contamination risks in flexible endoscopy.

REFERENCES

[1] C. L. Ofstead, M. R. Quick, J. E. Eiland, and S. J. Adams, "A glimpse at the true cost of reprocessing endoscopes," *International Association of Healthcare Central Service Material Management*, 2017.

[2] M. W. Gifari, H. Naghibi, S. Stramigioli, and M. Abayazid, "A review on recent advances in soft surgical robots for endoscopic applications," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 15, no. 5, p. e2010, 2019.

[3] M. Chauhan, J. H. Chandler, A. Jha, V. Subramaniam, K. L. Obstein, and P. Valdastri, "An origami-based soft robotic actuator for upper gastrointestinal endoscopic applications," *Frontiers in Robotics and AI*, vol. 8, p. 664720, 2021.

[4] J. H. Chandler, M. Chauhan, N. Garbin, K. L. Obstein, and P. Valdastri, "Parallel helix actuators for soft robotic applications," *Frontiers in Robotics and AI*, vol. 7, p. 119, 2020.

[5] N. Garbin, A. P. Mamunes, D. Sohn, R. W. Hawkins, P. Valdastri, and K. L. Obstein, "Evaluation of a novel low-cost disposable endoscope for visual assessment of the esophagus and stomach in an ex-vivo phantom model," *Endoscopy international open*, vol. 7, no. 09, pp. E1175–E1183, 2019.

Towards a Sensorised Balloon for Lumen and Polyp Detection using Electrical Impedance Tomography

Jackson Williams, Pietro Valdastri, James Avery, and James H. Chandler

STORM Lab, School of Electronic and Electrical Engineering, University of Leeds

INTRODUCTION

Traditional endoscopes rely primarily on camera-based imaging for diagnosis, however vision-based methods remain limited by poor illumination, occlusion, restricted depth perception, and an inability to detect subsurface or non-polypoid lesions [1]. As a result, clinically relevant information regarding tissue structure and mechanical properties may be missed when relying on visual feedback alone. To address these limitations, contact-based sensing approaches have been proposed to complement visual modalities by providing information derived from direct interaction with surrounding tissue. In clinical practice, tissue palpation is widely used to assess tissue type and identify abnormalities based on mechanical properties such as stiffness and compliance. Work has been completed to apply electrical impedance tomography (EIT) to create distributed strain sensors, enabling the measurement of deformation through changes in impedance [2]. This provides a means of capturing contact-induced variations in shape, offering information on geometry and material properties that are not accessible through vision alone.

In this work, a sensorised balloon (Figure 1) is developed for integration at the tip of an endoscope. The device incorporates longitudinal saline-filled channels, allowing changes in impedance to be measured during inflation. As the balloon expands within a lumen, its deformation is influenced by contact with surrounding structures, producing measurable changes in the electrical impedance (EI) response based on the cross-section of the channels. This enables the detection of lumen boundaries and sensitivity to localised geometrical perturbations. The proposed device has the potential to enhance minimally invasive procedures by improving the reliability of abnormality detection, supporting surveillance following polyp removal, and enabling more precise and informed interventions.

MATERIALS AND METHODS

The balloon is 15 mm in diameter and 25 mm in length (Figure 1) and is comprised of two parts: a balloon membrane with ten evenly spaced saline channels, and an inner 7-lumen catheter (Nusil MED-4880 silicone).

A three part 3D printed mould (Form4, Formlabs) with ten 0.8 mm nitinol rods was used to cast the balloon membrane in degassed EcoFlex 00-50 (SmoothOn) silicone. The membrane was demoulded and bonded to the multi-lumen section with Sil-poxy, producing a doughnut-shaped balloon with a single inflation inlet and ten channels running longitudinally. A 1:1 mix of 0.8 % saline solution and glycerine was used as the conductivity medium in the channels with added red food colouring for visualisation during testing. This combination was selected due to the conductive properties of saline and the ability of glycerine to reduce bubble formation [3].

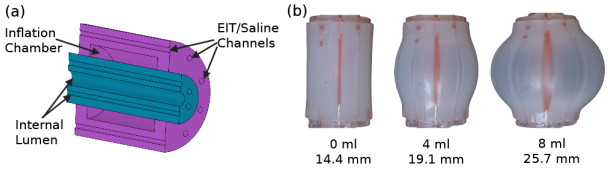


Fig. 1: (a) EI balloon cross section and (b) diameter at 0 ml, 4 ml and 8 ml syringe displacement.

The system comprises an Arduino Uno R4 microcontroller (MCU) interfaced with a Gecko G201X stepper motor driver, powered by a 105 W supply (35 V, 3 A). Actuation is provided by a NEMA 23 stepper motor (MOTECH MT-2303HS200AW-C), coupled to a linear rail and syringe carriage for volumetric control of balloon inflation using a 12 ml syringe. Pressure measurements are obtained using an Adafruit MPRLS ported pressure sensor. Imaging for diameter analysis in ImageJ is performed using a Basler camera (acA2040-120uc, Balser), which is externally triggered by the MCU via an I/O connection (Opto-I/O Cable, HRS 6p/open, P, Basler). Electrical impedance measurements are acquired using an impedance spectroscopy system (Eliko Quadra), connected to the EI channel via a 0.3 mm silver wire. Silver electrodes were selected due to their low electrode–electrolyte contact impedance in saline, providing reduced noise and improved measurement stability during acquisiton.

RESULTS

To characterise the EI response, three experimental conditions were evaluated: (i) unconstrained inflation, (ii) inflation within a 23 mm diameter rigid lumen, and (iii) inflation within the same lumen with a 3 mm hemispherical inclusion, aligned with the saline channel. Following each inflation step, a 10 s settling period was allowed.

For the baseline test, the balloon was inflated in free space, to determine the relationship between the syringe displacement and balloon diameter. As shown in Figure 1(a), the balloon diameter increases nonlinearly with syringe displacement, showing an initially rapid expansion from 14.4 mm at 0 ml, followed by a gradual reduction in growth rate, reaching 25.7 mm at 8 ml. Using linear interpolation, contact with the lumen and inclusion is

estimated to occur at 6.2 ml and 5.3 ml, respectively.

The internal balloon pressure, measured in 0.5 ml increments, follows a broadly linear increase across all test conditions from 0 kPa at 0 ml to 17 kPa at 2.5 ml, after which the rate of increase reduces and transitions towards a plateau (Figure 2(b)). This trend is consistent across the tests until approximately 5 ml (26.2 kPa), indicating unconstrained behaviour. Above this point, the the 'Lumen-constrained with inclusion' pressure diverges, increasing at 5.5 ml which aligns with the predicted inclusion contact point (5.3 ml) identified from Figure 2(a). In the lumen-only case, a similar increase in pressure is observed from approximately 6.5 ml, closely relating to the predicted lumen contact point (6.2 ml). By 8 ml, both constrained cases reach comparable pressures, approximately 34kPa, whereas the unconstrained case remains relatively flat. These results indicate that contact with the inclusion and lumen produces measurable internal pressure increases.

The general electrical impedance (EI) response shows a gradual, nonlinear increase with syringe displacement across all conditions, as seen in Figure 2(c). Starting from 3.5 kΩ at 0 ml, the impedance remains stable until 2 ml, after which it increases at a progressively higher rate, reaching 34.4 kΩ by 5 ml. In the inclusion case, a sharp increase in impedance occurs at 5.5 ml, rising to 189.6 kΩ; this represents an increase an order of a magnitude larger than the other results, which have a comparatively modest increase to 38.8 kΩ. At 6 ml, the EI channel in the inclusion case becomes effectively occluded, resulting in a loss of conductivity and measured impedance exceeding 4 MΩ. After the lumen contact point (6.2 ml), the lumen-constrained impedance increases more quickly to 86.4 kΩ at 8 ml compared to 76.6 kΩ for the unconstrained case.

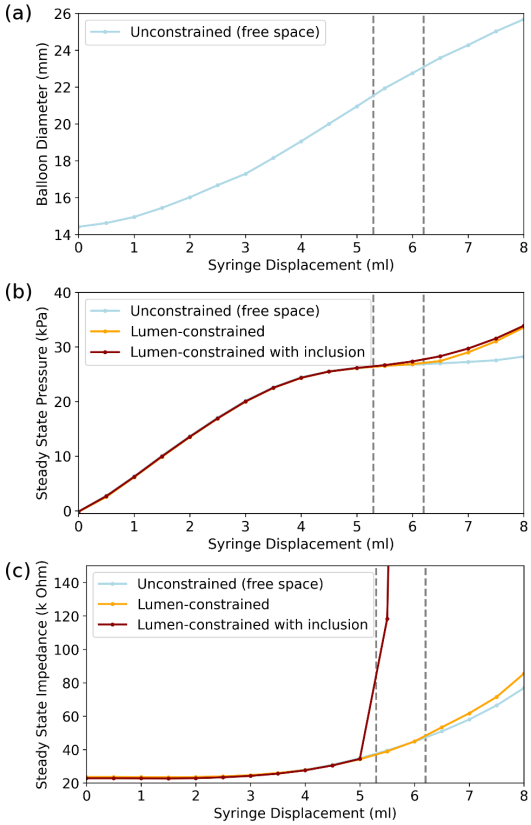


Fig. 2: (a) Balloon diameter as a function of syringe displacement. Plots showing (b) the internal balloon pressure and (c) the EI channel impedance as a function of syringe displacement. Vertical dashed lines at 5.3 ml and 6.2 ml indicate the estimated points of contact with the hemispherical inclusion and lumen boundary, respectively.

CONCLUSIONS AND DISCUSSION

The results demonstrate that both localised and distributed contact produce distinct and measurable changes in the mechanical and electrical response of the sensing balloon. In particular, deviations from the free-space behaviour in both pressure and impedance occur at displacement values closely aligned with the predicted contact points. This confirms that contact events can be inferred indirectly through changes in the inflation response.

The pressure measurements capture the global mechanical behaviour of the system, with increases observed once expansion becomes constrained by either the lumen or the inclusion. However, it does not distinguish between different types of contact. In contrast, the electrical impedance (EI) response shows greater sensitivity to localised deformation of the saline-filled channel. The presence of the hemispherical inclusion produces a rapid increase in impedance, exceeding one order of magnitude relative to the unconstrained case, followed by complete channel occlusion. This behaviour is markedly different from the more gradual impedance increase observed when the balloon contacts the lumen boundary, which reflects distributed constraint rather than localised obstruction.

This demonstrates that EI sensing provides complementary information to pressure measurements, which could enhance the detection of small or localised features, such as inclusions or lesions, that may not be easily distinguishable through global mechanical measurements alone. Further testing with compliant lumen and inclusions, which reflect tissue stiffness, are required to determine the efficacy of this method in a clinical setting, however these findings highlight the potential of the proposed approach for applications in endoscopic sensing.

Future work will build on this by extending from single-channel impedance measurement toward multi-channel systems and developing interconnected channel architectures capable of spatial reconstruction. This could allow radial localisation of contact and reconstruction of lumen shape.

REFERENCES

[1] Q. Wang, E. Chen, Y. Cai, C. Chen, W. Jin, Z. Zheng, Y. Jin, Y. Chen, X. Zhang, and Q. Li, "Preoperative endoscopic localization of colorectal cancer and tracing lymph nodes by using carbon nanoparticles in laparoscopy," *World Journal of Surgical Oncology*, vol. 14, no. 1, p. 231, Dec. 2016.

[2] A. Alian, J. Avery, and G. Mylonas, "Tissue palpation in endoscopy using EIT and soft actuators," *Frontiers in Robotics and AI*, vol. 11, p. 1372936, Aug. 2024.

[3] Y.-N. Cheung, Y. Zhu, C.-H. Cheng, C. Chao, and W. W.-F. Leung, "A novel fluidic strain sensor for large strain measurement," *Sensors and Actuators A: Physical*, vol. 147, no. 2, pp. 401–408, Oct. 2008.

Proximal Sensing of Instrument Rotation in Endovascular Procedures

Maria Beatriz da Costa Lopes, Wim-Alexander Beckers, Emmanuel Vander Poorten, and Beatriz Farola Barata

Robot-Assisted Surgery Group, Department of Mechanical Engineering, KU Leuven, Belgium

INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death globally [1]. Minimally invasive procedures to treat CVDs involve the use of guidewires and catheters to navigate the vasculature to a target site. This is achieved by manually manipulating instruments at their proximal end while typically relying on fluoroscopy for intra-procedural visualisation [2]. Yet, fluoroscopy has several drawbacks, including exposure to ionising radiation [3].

To address these drawbacks, there has been extensive research on how to incorporate sensors in endovascular instruments to enable their fluoroscopy-free visualisation [4]. Fibre optical shape sensing (FOSS) enables real-time three-dimensional (3D) shape visualisation of endovascular instruments by analysing changes in the wavelength spectra of light reflected through embedded optical fibres [5]. However, this technology alone does not provide information on the pose of the instrument with respect to a global reference frame. Al-Ahmad *et al.* used a 3D reference geometry to localise a catheter with respect to a global reference frame, allowing the direct estimation of the catheter insertion length with respect to a fixed point; however, this method alone does not track axial rotation [6].

Accurately computing the instrument rotation is critical for navigation in endovascular interventions, especially for catheters and guidewires with a pre-bent tip. When proximally rotating them, the clinician changes their tip's curve plane, allowing entry into different branches of the vasculature from the same position and directing the path to the desired location.

This work proposes a method to compute the real-time proximal rotation of an endovascular instrument with an embedded optical fibre relative to a reference curvature.

MATERIALS AND METHODS

Similarly to the work of Al-Ahmad *et al*. [6], this work uses a 3D reference shape with curvature, $\boldsymbol{\kappa}_{ref}$, and compares it with the curvature profile of a fibre-embedded instrument, $\boldsymbol{\kappa}$. In this work, the instrument is referred to as a stylet and consists of a nitinol tube with an embedded fibre Bragg grating (FBG) multi-core optical fibre. Both components are attached at the proximal connector, while the fibre remains unconstrained within the lumen of the tube. This design is important because fibre torsion cannot be measured, which would lead to incorrect shape reconstructions when the stylet is rotated. Consequently, when the stylet is rotated while inside the reference shape, the angles of the bending plane in the local fibre cross-section frame, computed from the strains in each fibre core, might not reflect the rotation of the stylet itself.

To address this limitation, a stepper motor (NEMA 17, 123 3D B.V., Almere, Netherlands) is introduced to rotate the proximal end of the stylet (Figure 1(A)), minimising the torsion effect, and decreasing the error between the rotation value derived from the bending angle measured by the FOSS system and the rotation applied to the stylet. The FOSS system used is an FBG-scan optical interrogator (FBGS International NV, Geel, Belgium).

The rotation of the stylet, φ, is computed as

$$\varphi = atan2(\sin(\Delta\theta), \cos(\Delta\theta)) \quad \textbf{(1)}$$

where $\Delta\theta$ corresponds to the difference between the initial and current bending angle at the point of maximum curvature in $\boldsymbol{\kappa}_{ref}$.

The stepper motor rotation, β, is defined as

$$\beta = min[\alpha\varphi, S_{max}] \quad \textbf{(2)}$$

with α being a scaling factor and S_{max} a maximum step for the rotation of the motor to ensure stable actuation, with these parameters being dependent on the FBG interrogation speed. The actuation of the motor operated synchronously with the FBG interrogator at a scan rate of 50 Hz.

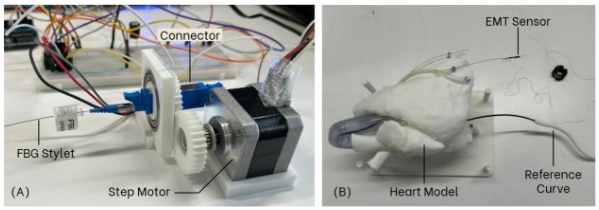


Figure 1. Experimental setup **(A)** of the motor-assisted rotation interface and artificial heart model **(B)** used for evaluation.

EXPERIMENTS AND RESULTS

Three experiments were conducted, and the obtained values for the rotation were compared with ground truth (GT) measurements from an electromagnetic (EM) tracking sensor (Northern Digital Inc., Waterloo, Ontario, Canada) attached to the tip of the stylet. The GT rotation angle, φ_{EMT}, with sub-degree orientation accuracy, was extracted from the EM sensor orientation data, provided as 3-DOF quaternions, q, using [7]:

$$\varphi_{EMT} = atan2\left(2(q_w q_z + q_x q_y), 1 - 2(q_y^2 + q_z^2)\right) \quad \textbf{(3)}$$

The first experiment consisted of the baseline performance assessment of the FBG system, without the stepper motor. In this case, the effect of torsion in the stylet's external tube is most evident. In the second experiment, motor-assisted stylet rotation was evaluated

This work was supported in part by VLAIO (Flanders Innovation & Entrepreneurship) under grant HBC.2023.0728 (ENLIGHT project) and in part by the LumiSense Project, an internally funded C3 Project of KU Leuven under grant 3E240097.

with the stylet in free space. In the third experiment, the system was tested when the stylet was inserted in an artificial heart model to analyse its behaviour in a more realistic environment. For each experiment, ten trials were performed. Table 1 summarises performance metrics, including the error defined by the difference between the GT and computed rotations, and Pearson correlation (PC), across all experimental setups.

Table 1. Error and PC (median [interquartile range]) of the rotation tracking system, against the GT, across all ten trials for each experimental setup: baseline, motor-assisted free space, and motor-assisted artificial heart model (MA: motor assisted).

	Error (°)	PC
Baseline	68.24 [28.48, 109.75]	0.334 [0.282, 0.363]
Free Space (MA)	50.60 [19.49, 87.25]	0.706 [0.386, 0.909]
Heart Model (MA)	15.80 [5.15, 35.40]	0.562 [0.227, 0.612]

Figure 2 shows the stylet rotation over time in free space with the stepper motor.

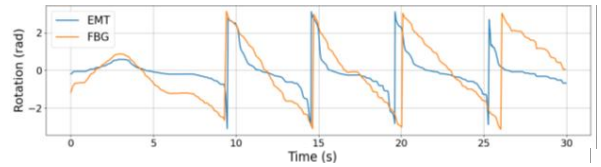


Figure 2. Rotation angle of the stylet over time in free space with motor assistance (blue: ground truth, orange: FBG measurement).

Figure 3(A) shows the results of the experiment testing the stylet with the stepper motor, inserted into the artificial heart model (Figure 1(B)). Figure 3(B) presents the absolute angular error distribution between motor-assisted rotation in the artificial heart model and GT, across all trials.

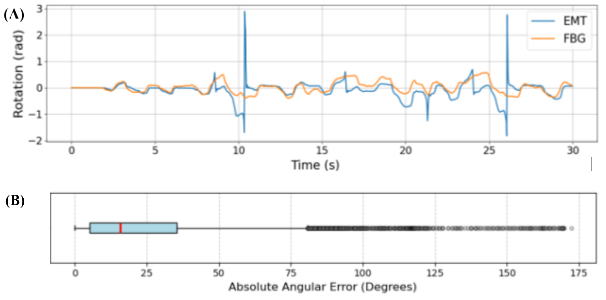


Figure 3. (A) Rotation angle of the stylet over time in the artificial heart model with motor assistance (blue: ground truth, orange: FBG measurement). **(B)** Absolute angular error distribution between motor-assisted rotation in the artificial heart model and GT.

DISCUSSION AND CONCLUSION

Accurate tracking of the endovascular instruments' rotation is critical during their navigation in interventions. This work presents a method capable of estimating this rotation in a stylet in real time by pairing FBG with a stepper motor.

For both motor-assisted experimental setups, results show agreement between the GT and the computed rotations, and that motor assistance not only reduces the error but also increases PC when compared with the baseline experiment. In these motor-assisted setups, the verified errors are likely still due to torsion and friction effects between the stylet's outer tube and the embedded optical fibre, which become more evident during rapid rotational changes.

Moreover, during stylet navigation inside the artificial heart model, rapid rotational changes occur due to stick-slip phenomena caused by friction between the stylet and the model. These phenomena can be seen as sharp peaks (see Figure 3 (A)) in the GT EM measurements, but they are absent in the proposed FBG-based approach, prompting the verified PC decrease in the artificial heart model setup compared to the free space setup. Nevertheless, the overall error is substantially reduced due to the more constrained rotational motion of the stylet inside the model.

The proposed FBG-based rotation tracking approach demonstrates potential, but its current limitations make it unsuitable as an independent clinical solution. Future work will focus on exploring multi-sensor fusion frameworks using the proposed approach as a complementary modality and miniaturising the actuation module to facilitate clinical integration into existing endovascular workflows.

REFERENCES

[1] World Health Organization, “Cardiovascular diseases,” https://www.who.int/health-topics/ cardiovascular-diseases#tab=tab_1, 2025, accessed: 2025-12-05.

[2] H. Sharei, T. Alderliesten, J. van den Dobbelsteen, J. Dankelman “Navigation of guidewires and catheters in the body during intervention procedures: a review of computer-based models,” Journal of Medical Imaging, vol. 5, no. 1, pp. 010 902–010 902, 2018.

[3] M. Roshanfar, M. Salimi, S.J. Jang, A.J. Sinusas, J. Kim, B. Mosadegh, “Emerging image-guided navigation techniques for cardiovascular interventions: A scoping review,” Bioengineering, vol. 12, no. 5, p. 488, 2025.

[4] B. Farola Barata, P.T. Tran, G. Borghesan, K. McCutcheon, D. Dall'Alba, P. Fiorini, J Vander Sloten “IVUS-based local vessel estimation for robotic intravascular navigation,” IEEE robotics and automation letters, vol. 6, no. 4, pp. 8102–8109, 2021.

[5] M. Jansen, A. Khandige, H. Kobeiter, E.J. Vonken, C. Hazenberg, J. van Herwaarden “Three-dimensional visualisation of endovascular guidewires and catheters based on laser light instead of fluoroscopy with fiber optic realshape technology: preclinical results,” European Journal of Vascular and Endovascular Surgery, vol. 60, no. 1, pp. 135–143, 2020.

[6] O. Al-Ahmad, M. Ourak, J. Van Roosbroeck, J. Vlekken, E. Vander Poorten “Real-time fiber-optic shape registration for catheter localization in endovascular procedures.” In Proceedings of the Joint Workshop on New Technologies for Computer/Robot Assisted Surgery (CRAS) 2023.

[7] Z. Yaniv, E. Wilson, D. Lindisch, K. Cleary “Electromagnetic tracking in the clinical environment,” Medical Physics, 36(3), 876–892, 2009.

Towards Combined Diffuse Reflectance and Electrical Impedance Spectroscopy for Colorectal Tissue Characterisation

L. Wachtel[1], N. Mandal[1], J. Williams[2], A. Yanez[2], J. Avery[2], and A. J. Thompson[1]

[1]*Helen Hamlyn Centre for Robotic Surgery, Institute of Global Health Innovation, Imperial College London*

[2]*School of Electronic and Electrical Engineering, University of Leeds*

INTRODUCTION

Colonoscopy is the current gold standard for detecting colorectal cancer (CRC) [1], but histopathological assessment of biopsy tissue is required for confirmation. This often results in the removal of benign or healthy tissue [2] and delays diagnosis.

Point-of-care tissue assessment is desired. This would facilitate *in vivo* diagnosis. The ideal method would detect both the structural and molecular/functional changes associated with cancer, as multimodal approaches can improve sensitivity and specificity [3].

We therefore propose a single probe integrating diffuse reflectance spectroscopy (DRS) and electrical impedance spectroscopy (EIS) - see Figure 1. Together, they provide access to complementary optical, electrical and microstructural tissue properties [3], [4]. Prior work has demonstrated a similar optical and electrical sensing combination for breast tissue assessment [5], thereby supporting our proposal's feasibility while leaving scope for novel application to CRC.

A multimodal probe's diagnostic value depends on the contrast mechanisms and the tissue volumes sampled by each modality. We therefore present a simulation framework to compare modality sensitivity distributions and explore the impact of probe geometry. This work then reports a prototype integrating a tetrapolar electrode array with a commercial DRS probe, laying the foundation for a future endoscope-compatible probe.

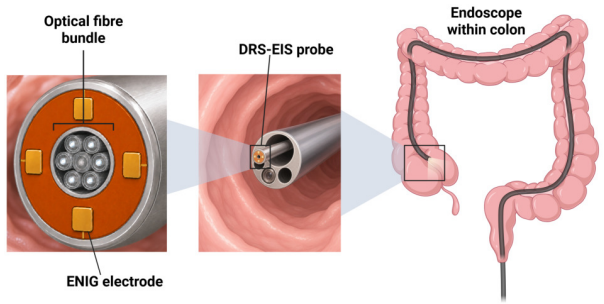


Fig. 1: Endoscopic DRS–EIS probe concept. Created with BioRender.

MATERIALS AND METHODS

BENCHTOP SYSTEM

The pilot system was assembled from off-the-shelf DRS and EIS hardware, integrated using bespoke mechanical mounts. The DRS equipment consisted of a miniature spectrometer (Flame, Ocean Optics), a 6.35 mm diameter fibre-optic probe (QR200-7-VIS-NIR, Ocean Optics), and a tungsten halogen light source providing illumination in the range 360-2400 nm (HL-2000-LL-FHSA, Ocean Optics). The probe possessed 7 fibres, each with a core diameter of 200 μm, in a 6-around-1 configuration. The EIS instrumentation comprised an impedance spectroscope (Quadra, Eliko) and a custom flexible printed circuit (FPC) with 4 embedded Electroless Nickel Immersion Gold (ENIG) electrodes, each 1 mm in diameter and arranged in a ring of 8.71 mm diameter. The FPC was affixed to the DRS probe with a 3D-printed mount (Figure 2).

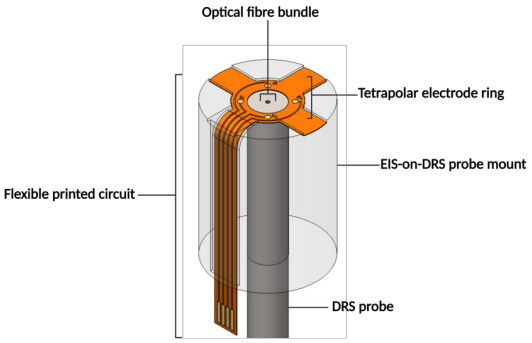


Fig. 2: SOLIDWORKS model of prototype. Created with BioRender.

SIMULATION

The EIS spatial sensitivity distribution was estimated from the Jacobian J of an Electrical Impedance Tomography and Diffuse Optical Tomography Reconstruction Software (EIDORS) finite element model, using the summed absolute Jacobian magnitude across simulated measurements. DRS sensitivity was estimated in MCmatlab using a Monte Carlo-derived photon measurement density function (PMDF), calculated from the product of source and detector fluence fields. For both, large homogeneous media were used to approximate semi-infinite simulation domains, thus avoiding their boundaries impacting simulation results. Sensitivity maps (Figure 3) were normalised to unit total sensitivity. Characteristic depth was defined as the depth measured from the tissue surface within which 50% of the total volume-integrated sensitivity was contained.

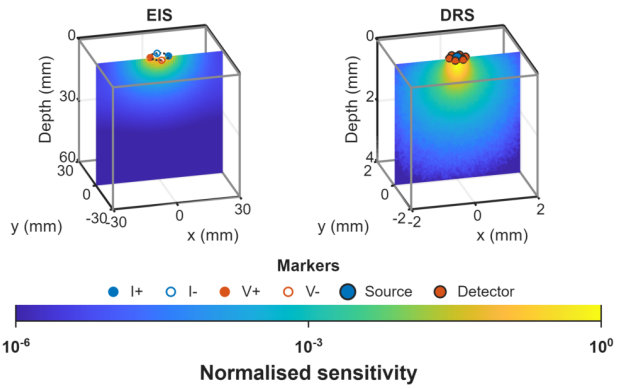


Fig. 3: Central plane sensitivity distributions for EIS and DRS, with electrodes and optical fibres denoted by circular markers.

IN VIVO FEASIBILITY DEMONSTRATION

Following simulation, the capacity of the benchtop system to simultaneously acquire optical and electrical spectra

from biological tissue was evaluated. The goal of this initial assessment was to prove prototype functionality. Validation on clinically meaningful samples is reserved for future work. Consequently, EIS and DRS measurements were concurrently acquired from the first author's index finger, which served as a convenient non-clinical *in vivo* target whose spectra are documented in the literature. The combined probe was clamped into place to prevent motion artefacts and complete probe-to-skin contact was verified. 5 measurements were taken from random locations on the finger pad then averaged. For DRS, 150 spectra were captured per location with all spectra undergoing dark correction and normalisation relative to a reflectance standard (USRS-99-010 AS-01158-060, Labsphere). DRS integration time was 60 ms (with 5 ms between acquisitions). For EIS, impedance across 15 frequencies was recorded (1 kHz - 349 kHz) for 30 s per location.

RESULTS AND DISCUSSION

SIMULATION

For DRS, characteristic depth was 0.78 mm, while for EIS, it was 1.21 mm. This indicates a shared superficial interrogation region, but also identifies that the current EIS configuration probes deeper into the tissue than DRS. These results demonstrate the framework's value for identifying the extent of sensitivity overlap and guiding future fibre and electrode geometry optimisation.

IN VIVO FEASIBILITY DEMONSTRATION

Figure 4 shows relative reflectance spectra from finger pad DRS measurements. Minima observed at approximately 425 nm, 548 nm and 577 nm indicate absorption by oxy- and deoxyhaemoglobin [6]. The spectra therefore exhibit known features of perfused skin.

Figure 5 presents impedance spectra from finger pad EIS measurements. A decrease in impedance magnitude with increasing frequency and a frequency-dependent phase response are observed. At low frequencies, current predominantly flows through the extracellular space, whereas at high frequencies, current can also flow through cells due to capacitive coupling [7]. The presence of an additional parallel conducting pathway at higher frequencies is likely responsible for the decrease in impedance magnitude. The deviation of impedance phase from 0° is attributed to capacitive effects of cell membranes at higher frequencies. The measured impedance data therefore reflects the known resistive and capacitive properties of skin, and also resembles spectra in the literature [8].

CONCLUSIONS

In summary, the results establish a framework for analysing multimodal (EIS/DRS) sensitivity distributions and the impact of probe geometry. Preliminary skin measurements using a prototype DRS-EIS probe demonstrate simultaneous acquisition of optical and electrical spectra from tissue. Simulations show that the current geometry provides a shared superficial interrogation region but that its spatial sensitivity overlap is imperfect, highlighting the necessity

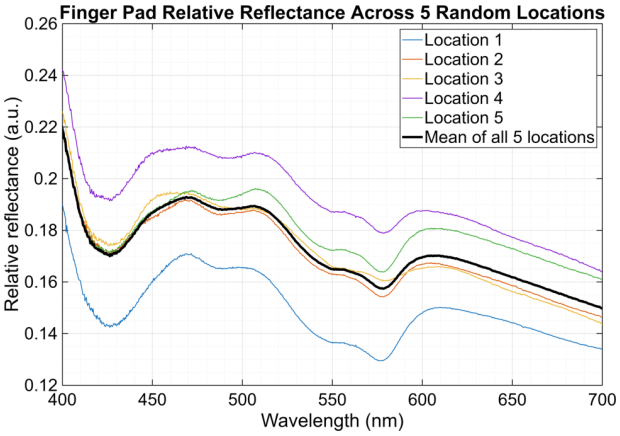


Fig. 4: Relative reflectance spectra, calculated with $R_{relative} = \frac{I_{\text{sample}} - I_{\text{dark,sample}}}{I_{\text{standard}} - I_{\text{dark,standard}}} \frac{t_{standard}}{t_{sample}}$. The thicker line is the mean of all 5 locations.

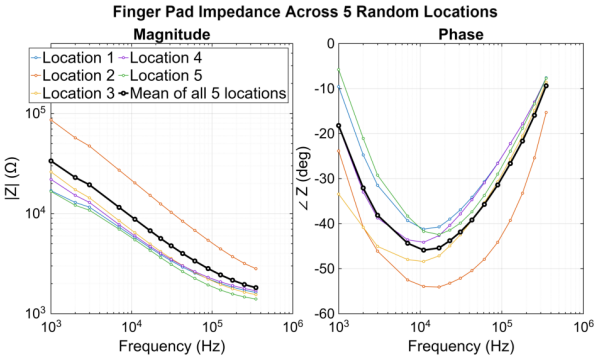


Fig. 5: Impedance spectra. (a) Magnitude against frequency. (b) Phase against frequency. The thicker line is the mean of all 5 locations.

of fibre and electrode placement optimisation. Further work is underway to evaluate the probe in *ex vivo* tissue.

REFERENCES

[1] Y. Zhang, Y. Wang, B. Zhang, P. Li, and Y. Zhao, "Methods and biomarkers for early detection, prediction, and diagnosis of colorectal cancer," *Biomedicine & Pharmacotherapy*, vol. 163, p. 114786, Jul. 2023.

[2] A. J. Thompson and G.-Z. Yang, "Tethered and Implantable Optical Sensors," in *Implantable Sensors and Systems: From Theory to Practice*, G.-Z. Yang, Ed. Springer International Publishing, 2018, pp. 439–505.

[3] V. Tuchin, J. Popp, and V. Zakharov, Eds., *Multimodal Optical Diagnostics of Cancer*, 1st ed. Springer International Publishing, Oct. 2020.

[4] M. Shokoufi, P. Grewal, and F. Golnaraghi, "Combining Two Breast Cancer Diagnosis Technologies: Diffuse Optical Tomography with EIS," *CMBES Proceedings*, vol. 37, May 2014.

[5] M. Shokoufi, "Multi-Modality Breast Cancer Assessment Tools Using Diffuse Optical and Electrical Impedance Spectroscopy," Ph.D. dissertation, Simon Fraser University, Nov. 2016.

[6] K. Setchfield, A. Gorman, A. H. R. W. Simpson, M. G. Somekh, and A. J. Wright, "Relevance and utility of the in-vivo and ex-vivo optical properties of the skin reported in the literature: a review [invited]," *Biomedical Optics Express*, vol. 14, no. 7, pp. 3555–3583, Jun. 2023.

[7] S. Ollmar, "Methods of information extraction from impedance spectra of biological tissue, in particular skin and oral mucosa—a critical review and suggestions for the future," *Bioelectrochemistry and Bioenergetics*, vol. 45, no. 2, pp. 157–160, 1998.

[8] S. M. Brunnquell, "Development of carbon electrodes for the measurement of skin impedance," Master's thesis, Universidade do Estado de Santa Catarina, Joinville, SC, Brazil, 2018.

Experimental Characterization of NURD in OCT Catheters Under Curvature and Rotational Speed Variations

Zeno Pogacnik[1], Giovanni Battista Regazzo[1], Daniel Corona Oliveira Costa[1], Oscar Caravaca-Mora[2], Emmanuel Vander Poorten[1] and Beatriz Farola Barata[1]

[1]*Robot-Assisted Surgery (RAS) Group, Department of Mechanical Engineering, KU Leuven, Belgium.*
[2]*Universidad Nacional de Costa Rica*

INTRODUCTION

Optical Coherence Tomography (OCT) is an imaging modality widely adopted in minimally invasive medical procedures. The integration of OCT into catheter systems enables real-time visualization of tissue morphology, assessment of device positioning, and evaluation of procedural outcomes without exposure to ionizing radiation [1].

OCT provides high acquisition speed and spatial resolution, typically ranging between 10 and 25 µm depending on the light source wavelength (1310 nm in the present system). The imaging principle relies on low-coherence interferometry. Light emitted through a rotating optical fiber with a lensed distal tip is backscattered by the sample and interfered with light reflected from a reference arm. The resulting signal provides depth-resolved information (A-line). A 2D cross-sectional image (B-scan) is obtained by sequentially acquiring A-lines during fiber rotation.

One of the major limitations affecting OCT image quality is Non-Uniform Rotational Distortion (NURD). NURD can be broadly classified into within-frame (WF-NURD) and between-frame (BF-NURD). WF-NURD refers to rotational non-linearities occurring within a single B-scan, producing local stretching or compression artifacts due to irregular angular velocity of the rotating optical element. These irregularities are generally associated with friction inside the catheter. BF-NURD instead appears across consecutive B-scans and manifests as circumferential jitter of image features. Such effects are commonly associated with rotational motor speed fluctuations and synchronization inaccuracies between image acquisition and fiber rotation.

Since eliminating NURD exclusively through hardware optimization is costly and impractical, several compensation algorithms have been proposed for both real-time and post-processing applications [2]. However, these approaches still rely on realistic assumptions regarding the temporal and spatial behaviour of the distortion. A better understanding of how NURD evolves under different conditions could therefore support both catheter design and more physically consistent compensation strategies.

The present work proposes an experimental methodology for characterizing the influence, on both WF-NURD and BF-NURD, of curvature radius and rotational speed, expressed as B-scan frame rate (fps). Previous studies reported that NURD in OCT catheters is weakly correlated with bending angle alone [3]; therefore, this variable was kept constant at 90° throughout the experiments.

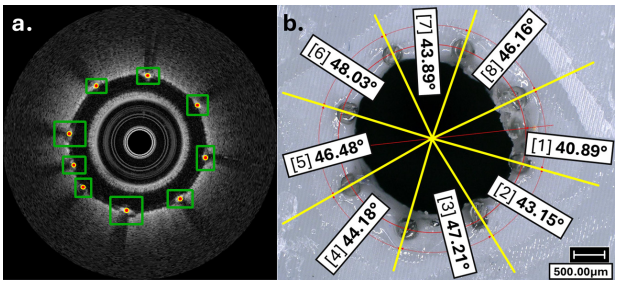


Fig. 1: a. Feature tracking in OCT frames. b. Microscope image of the octagonal phantom at outer cross-section.

MATERIALS AND METHODS

The experimental evaluation was conducted using a commercially available Dragonfly OpStar™ OCT catheter (Abbott, USA), designed for vascular imaging in vessels with diameters between 2.0 and 3.5 mm. The actuation and signal processing modules were developed as described in [4]. Distilled water was used inside the sheath to reduce refractive index discontinuities and improve image quality.

Controlled bending conditions were imposed by guiding the catheter through a PTFE sheath with an inner diameter of 1.2 mm inserted into 3D-printed supports. Curvature radii between 15 and 35 mm, with 5 mm increments, were investigated, covering a physiologically relevant range. Preliminary tests showed that a 10 mm radius introduced excessive friction, leading to stick slip rotational motion of the fiber and severe corruption of the reconstructed OCT images, and was therefore excluded. The frame rate was varied between 5 and 40 fps in steps of 5 fps.

To quantify the NURD, a custom phantom was developed consisting of eight steel wires with a diameter of 0.2 mm glued at the vertices of an octagon inscribable within a 3.5 mm circumference. A ninth wire was added as an absolute reference marker between two other wires. The phantom was manufactured using an MSLA 3D printer (HeyGears Reflex RS). Two perforated end plates were added to constrain and reproducibly center the catheter within the phantom during all acquisitions.

For each curvature and frame rate configuration, OCT images of the phantom were acquired. Because BF-NURD produces temporal oscillations of the wires positions, a feature tracking algorithm was used to follow the phantom markers across consecutive frames and retrieve their position (Fig. 1a).

The angular spacing between adjacent wires (angular sectors) was used as a measure of WF-NURD, since local

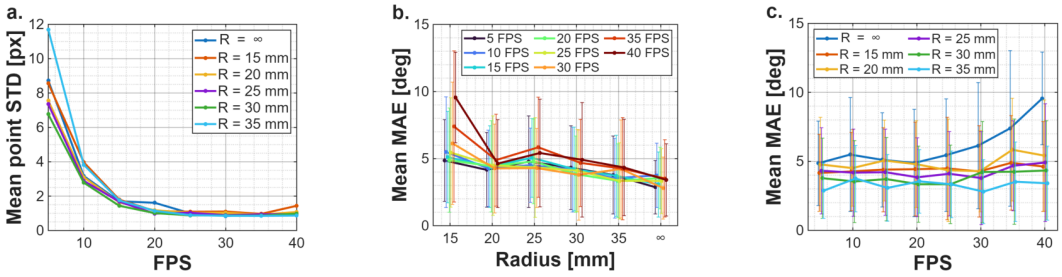


Fig. 2: a. BF-NURD vs frame rate. b. WF-NURD vs curvature radius. c. WF-NURD vs frame rate.

distortions result in compression and expansion of these angles. The angular positions of the phantom wires were measured on both outer faces of the phantom using a digital microscope (Keyence VHX-6000), and the average values were used as angular references (Fig. 1b). For each configuration a set number of frames were studied. For each frame, the mean absolute angular error (MAE) across the eight angular sectors was computed. The reported WF-NURD metric corresponds to the average of these MAE.

To quantify BF-NURD, the standard deviation of the tracked wire positions over time was computed for each acquisition, and its mean value was used as a stability metric.

RESULTS AND DISCUSSION

The BF-NURD analysis (Fig. 2a) showed that the positional instability decreases with increasing frame rate and reaches a stable minimum above approximately 20 fps (1200 rpm). This minimum corresponds to an average STD of ≈ 1 px and is largely independent of the curvature radius.

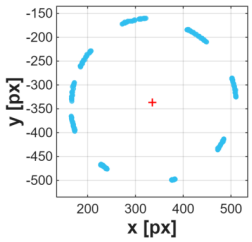


Fig. 3: Tracked points for 35 mm radius and 5 fps.

Figure 3 provides a visualization of the tracked point distribution. The points oscillate around a center of rotation that does not coincide with the catheter center. Similar behaviour was observed for all tested curvatures.

A possible explanation for this behaviour is the synchronization mechanism adopted by the OCT system. Since each B-scan is triggered by a photosensor detecting the passage of a shutter, all frames are initialized from the same angular reference position. However, fluctuations in instantaneous rotational velocity may progressively alter the circumferential placement of subsequent A-lines, resulting in the observed inter-frame oscillation of the features.

The WF-NURD results are reported in Figure 2b and 2c. As expected, smaller curvature radii generally increased the measured distortion, likely due to increased friction between the torque coil and the catheter sheath. However, the relationship is not strictly monotonic and several configurations deviated from the overall trend.

The frame rate showed a significant influence on WF-NURD only for the straight catheter configuration, suggesting that under bent configurations the effect of curvature-induced friction dominates over rotational-speed-dependent mechanisms. Furthermore, the observed quadratic trend in Figure 2c ($R^2 = 0.95$) may be partially attributed to viscous resistance generated by the distilled water inside the catheter.

Another relevant observation is the relatively large variability of the angular errors among each of the 8 sectors of the phantom. This indicates that the local deformations induced by WF-NURD are not uniformly distributed along the circumference. Since the variability remains approximately constant with increasing FPS, the angular sector deformation appears to remain relatively stable despite the temporal oscillations associated with BF-NURD.

The proposed methodology enabled the characterization of NURD behaviour under controlled bending and rotational conditions. Future work will focus on improving NURD assessment by considering a wider and more densely sampled range of curvature radii and frame rates, which may also help to better identify the transition between curvature-dependent WF-NURD regimes as a function of frame rate. Furthermore, WF-NURD assessment could be refined by restricting the analysis to frames acquired under rotational conditions closest to the target speed, identified either from frame-size consistency or from synchronized speed measurements.

Finally, the catheter used in this study had previously been employed during a clinical procedure, and its mechanical wear may therefore have influenced the reported results. Moreover, further optimization of the acquisition module for this specific catheter could likely improve image quality. Future investigations will apply the proposed methodology to custom-designed catheters.

Acknowledgement: This work was supported by by VLAIO (Flanders Innovation & Entrepreneurship) under grant HBC.2023.0728 (ENLIGHT project), by the LumiSense Project, an internally funded C3 Project of KU Leuven under grant 3E240097, and by the flemish research foundation (FWO, 1S35525N.

REFERENCES

[1] A. Y. Tarman, Ahmed *et al.*, "Smart catheters for diagnosis, monitoring, and therapy," *Advanced Healthcare Materials*, vol. 15, 2026.

[2] O. C.-M. Guiqiu Liao *et al.*, "Distortion and instability compensation with deep learning for rotational scanning endoscopic optical coherence tomography," *Medical Image Analysis*, vol. 77, 2022.

[3] Y. Kawase, Suzuki *et al.*, "Comparison of nonuniform rotational distortion between mechanical IVUS and OCT using a phantom model," *Ultrasound in Medicine and Biology*, vol. 33, 2007.

[4] O. M. Caravaca Mora, "Development of a novel method using optical coherence tomography (oct) for guidance of robotized interventional endoscopy," Ph.D. dissertation, Université de Strasbourg, 2020, nNT: 2020STRAD004. tel-03510154.

Shape control of a concentric tube continuum robot for surface ECoG brain implant insertion in a highly constrained environment

Ambroise Ulrich[1,2,3], Edwige Bano[2], Blaise Yvert[3], and M. Taha Chikhaoui[1]

[1] *Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, 38000 Grenoble, France*

[2] *Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, UMR 5130, Grenoble INP, CROMA, 38000 Grenoble, France*

[3] *Univ. Grenoble Alpes, Inserm, UMR 1216, GIN, 38000 Grenoble, France*

INTRODUCTION

Recent advances in neural implants have enabled patients with severe paralysis to recover functions such as locomotion and speech by decoding and transmitting brain signals to prostheses [1], [2]. These applications require coverage of large cortical areas involved in motor control and speech production, leading to wide craniotomies and therefore increased surgical risk.

To reduce invasiveness, we propose a continuum robotic system for deploying soft rectangular implants across multiple cortical sites through a small craniotomy. The device operates within the confined space between the skull and brain while preserving tissue integrity, with the objective of maintaining a distance from brain surface below 1mm. A Concentric Tube Continuum Robot (CTCR) is selected for its flexibility and suitability for navigation in constrained environments [3]. It consists of three pre-curved, nested tubes, each actuated in rotation and translation, cf. Fig. 1.

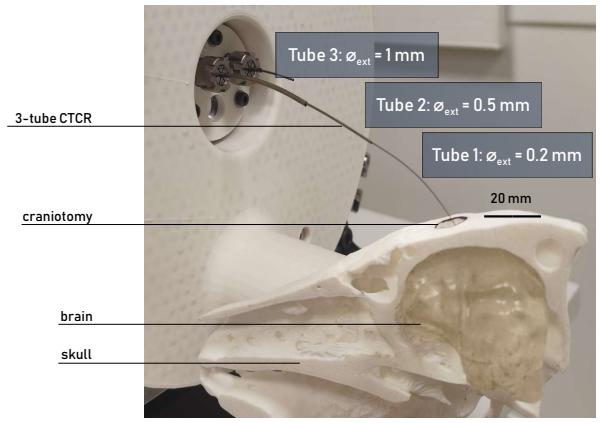


Fig. 1: Inserting a 3-tube CTCR inside a phantom of minipig skull through a 11mm diameter craniotomy.

The objective of this work is to compute a trajectory for a CTCR to reach targets while following the curvature of the brain. For preclinical validation, the system is designed to insert implants on minipig brain, targeting cortical regions involved in vocalization. The craniotomy is located at the top of the skull and limited to a diameter of 11mm, following clinical recommendations.

METHODS

Anatomical meshes of the brain and skull of a minipg are used for simulation. Cortical regions of interest are identified, from which two targets are defined, each specifying a position and orientation (pose) required for implant deployment. The desired trajectory is designed to guide the robot to the targets while avoiding contact with brain tissue. In [4], a multi-objective Particle Swarm Optimization (PSO) algorithm is described, resulting in a robot design with final configurations reaching target poses 1 and 2 while avoiding tissue interaction as presented on Fig. 2. These final robot configurations serve as reference shapes for the shape control method, with each target trajectory discretized into successive waypoints tracked by the control algorithm.

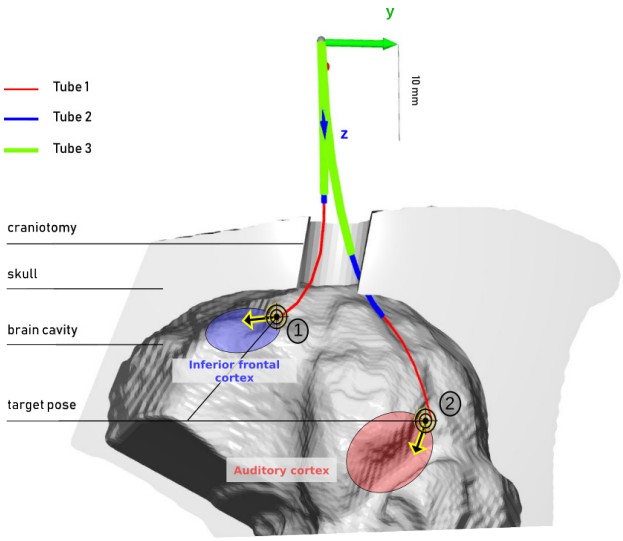


Fig. 2: Robot configurations reaching two target poses over relevant minipig cortex.

The robot control is based on a Follow-The-Leader (FTL) strategy, in which the robot progressively describes its final shape, minimizing swept volume and avoiding contact with tissue during insertion. The FTL trajectory is computed using a Generalized Damped Least-Squares (GDLS) formulation [3]. The controller guides the robot through the prescribed waypoints by simultaneously satisfying three control objectives: tip position control, shape control, and tube extension regularization.

Tip control drives the robot tip toward the successive

waypoints, while tube extension regularization promotes a stable deployment by preventing excessive or unbalanced tube motions. Shape control enforces FTL behavior through a set of sub-constraints, each attracting a point along the robot centerline toward its closest point on the reference shape, following the approach proposed in [5]. The combination of these objectives enables accurate waypoint tracking while maintaining a stable configuration and preserving the desired FTL behavior throughout the deployment.

RESULTS

The selected reference shapes reach Target 1 with a maximum distance to the brain surface of 1.29 mm and Target 2 with a maximum distance of 1.01 mm. The corresponding reference shapes are discretized into waypoints tracked using the GDLS algorithm. Computation time increases proportionally with the number of waypoints. Since no performance improvement is observed beyond, 30 WP are retained (cf. Fig. 3e). This control strategy produces a sequence of actuations that enables the robot to follow the reference shape while approximating a FTL deployment.

The error between the robot and the reference shape is computed by first interpolating the reference trajectory from its waypoints, and then evaluating, for each point along the robot, the minimum distance to this interpolated shape.

For Target 1, the mean shape error during deployment is 0.26 mm, with a standard deviation of 0.2 mm and a maximum error of 1.0 mm. For Target 2, the mean error is 0.48 mm, with a standard deviation of 0.4 mm and a maximum error of 2.5 mm. The final tip error is 0.4 mm for Target 1 and 1.1 mm for Target 2. These correspond to relative errors of 1.1% for a deployed length of 35.4 mm and 2.0% for a deployed length of 53.7 mm, respectively, as illustrated in Fig. 3. Without tip orientation control, the final tip orientation error reaches 2.3° and 6.2° for Targets 1 and 2, respectively.

Figs. 3c and 3d compare the error between the robot shape and the reference shape for controllers without and with shape control. The results show a significant improvement when shape control is enabled, with the maximum shape error decreasing from 4mm to 2.5mm for Target 2. In contrast, no noticeable improvement is observed for Target 1, which is located closer to the robot base.

CONCLUSIONS AND DISCUSSION

We present a CTCR control framework to reach cortical targets while maintaining proximity to the brain surface using a single robot design. While effective, the current FTL strategy remains imperfect; during deployment, the distance between the robot and the reference trajectory reaches up to 2.5 mm, exceeding the targeted 1 mm threshold.

Future work will focus on reducing the error between the robot and the reference shape. Possible directions include better fine tuning the control weights, or directly constraining the robot to follow the brain surface rather than the reference trajectory, thereby providing higher flexibility to the control algorithm during deployment.

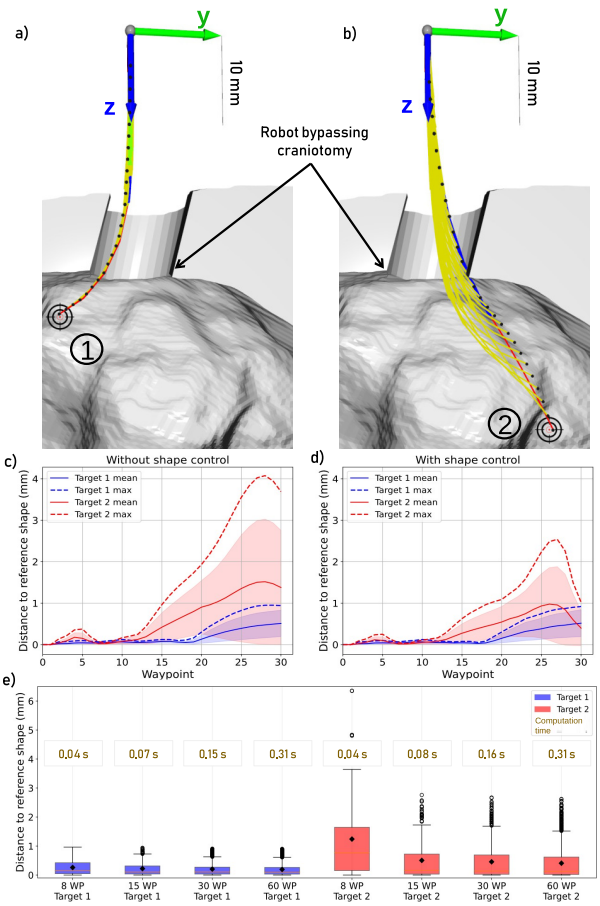


Fig. 3: Control results: a), b) Representation of swept volume by FTL, respectively for Targets 1 and 2 with in yellow the intermediate configurations during robot deployment and in colors the final shape of the robot; c), d) Error between robot centerline and reference shape at each control step respectively without and with shape control; e) Boxplot comparing shape error for different number of waypoints in shape control scenario.

Once the framework is fully validated, the next step toward preclinical applicability is experimental insertion of a CTCR prototype into a phantom minipig skull toward predefined targets. This will allow assessment of the method's ability to design and control the robot to reach targets while minimizing the pressure applied to the brain.

REFERENCES

[1] S. Duraivel *et al.*, "High-resolution neural recordings improve the accuracy of speech decoding," *Nature Comm.*, 14(1): 6938, 2023.

[2] H. Lorach, A. Galvez, V. Spagnolo *et al.*, "Walking naturally after spinal cord injury using a brain–spine interface," *Nature*, 618(7963): 126-133, 2023.

[3] J. Burgner *et al.*, "A telerobotic system for transnasal surgery," *IEEE/ASME Trans. Mechatron.*, 19(3): 669-1006, 2014.

[4] A. Ulrich *et al.*, "Design optimization of a concentric tube continuum robot for surface ecog brain implant insertion in a highly constrained environment," *CRAS*, pp. 115–116, 2025.

[5] Y. Xu *et al.*, "Shape control of concentric tube robots via approximate follow-the-leader motion," *IEEE Robot. Autom. Lett.*, 9(8): 7198-7205, 2024.

Acknowledgments: This work is supported by grants ANR-15-IDEX-02, ANR-24-CE33-4611, and ANR-23-IACL-0006.

Magnetic Force Sensing in Ophthalmic Surgical Robot Forceps: Design, Integration, and Validation

Charles DeLorey[1], Benjamin Jackson[1], Mojtaba Esfandiari[2], Pietro Valdastri[1], Iulian Iordachita[2], and Dominic Jones[1]

[1]*School of Electronic and Electrical Engineering, University of Leeds, UK*
[2]*Whiting School of Engineering, Johns Hopkins University, USA*

INTRODUCTION

Ophthalmic surgery demands exceptional precision and control to avoid tissue damage in delicate procedures such as cataract removal, retinal repair, and corneal transplantation. Robotic surgical systems can enhance consistency and reduce surgeon fatigue; however, they often lack adequate force feedback necessary for safe manipulation of fragile ocular tissues [1].

Force sensing in the gripper is important for preventing excessive grasping forces, providing haptic feedback, and ensuring safe tissue manipulation [2]. Conventional force sensors are often too large or fragile for surgical forceps integration. While fiber Bragg grating (FBG) sensors enable measurement along the instrument length [3], they do not provide direct force sensing of tissue in the forcep grip. Similar prior work has demonstrated force sensing micro-forceps for robot-assisted retinal surgery, however detection of in-jaw contact forces has still not been achieved [4]. To address the need for direct grip force measurement, we designed a compact 1 mm × 1 mm × 0.7 mm magnetic force sensor specifically to measure grip force at the tissue contact point.

MATERIALS AND METHODS

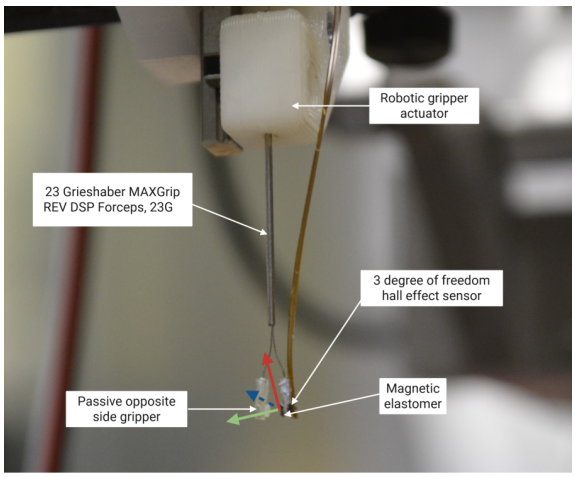
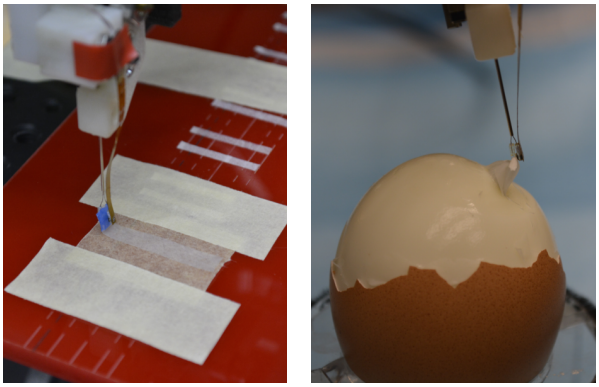


Fig. 1: Integrated forceps with magnetic force sensor showing gripper design and miniaturized form factor. Sensor coordinate frame: X (blue), Y (red), Z (green).

We designed a force sensor using Hall effect sensing with magnetic elastomers. The sensing principle is that when force is applied to the sensor, the magnetic elastomer deforms, causing a measurable change in the local magnetic field strength. Hall effect sensors detect this field variation, which is calibrated to the applied force.

Sensor calibration was achieved using a 6 degree-of-freedom (DOF) force/torque transducer (ATI Industrial Automation, Apex, NC, USA) attached to a parallel robot platform (Symétrie, Nîmes, France). The robot applied controlled loads to the ophthalmic sensor face, recording raw sensor outputs and ground-truth force measurements from the ATI sensor. 50 indentation points were performed on the elastomer surface. We trained a neural network (2 × 64 hidden layers) to map Hall sensor readings to 3D contact forces, calibrating the sensor across a measurement range of 0–500 mN for gripping force and 0–150 mN for pulling/shearing forces. The sensor components were integrated into a minimally invasive forceps design (23G Alcon Grieshaber ophthalmic forceps: 30 mm x 0.6 mm⌀ shaft, 0.2 mm jaw width, 0.1 mm thickness) with the 1 mm × 1 mm × 0.7 mm form factor (Figure 1), allowing integration with the ophthalmic surgical robot arm [5].



(a) Egg peeling trial setup with integrated magnetic force sensor in forceps.

(b) Close-up view of forceps gripping egg membrane during peeling trial.

Fig. 2: Validation procedures including tape peeling trials and biological validation through egg peeling experiments.

Tape peeling trials (n=10) were conducted to establish baseline sensor performance with controlled kinematics. Pre-programmed robot movement moved the forceps 15 mm at a 45° angle to lift and pull the tape. Biological

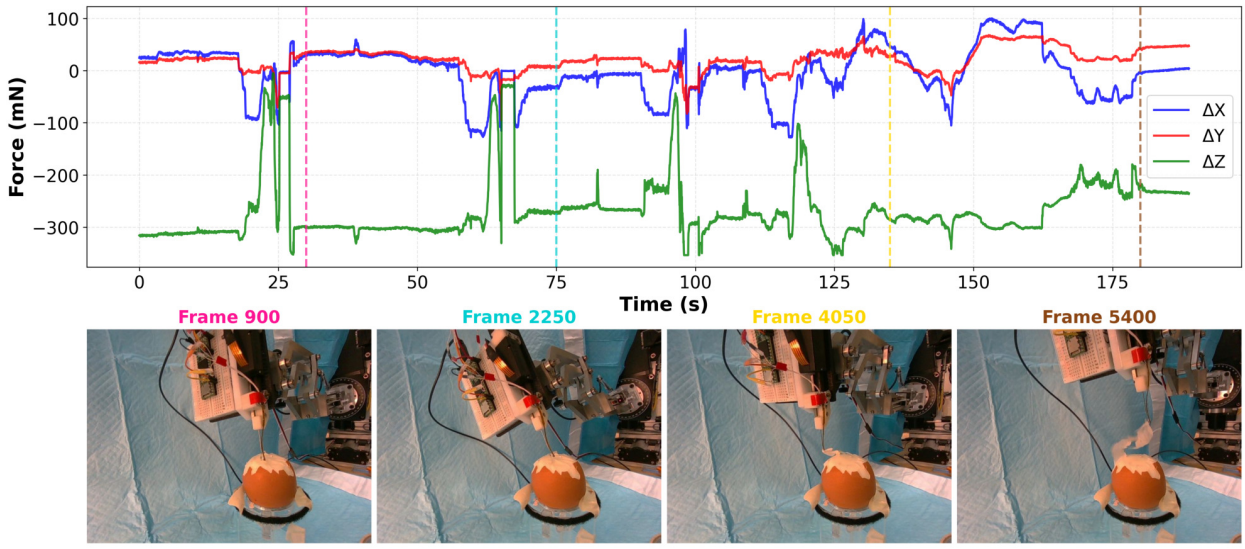


Fig. 3: Force measurements during egg peeling trials showing grip force feedback across the three degrees of freedom.

validation was performed via human teleoperated egg peeling experiments (n=5).

RESULTS

Tape Peeling Trial: Across 10 controlled trials, the minimum and maximum measured forces remained within 12% variance for each axis. The sensor successfully recorded a maximum force of 140 mN in the Z-axis and 18 mN in the X-axis, demonstrating highly consistent performance during controlled manipulation tasks.

Egg Peeling Experiments: During the biological validation trials, operators successfully manipulated the delicate egg membrane to perform complete peels in all 5 trials. Throughout the task, the sensor capably recorded dynamic forces (see Figure 3), capturing a maximum force of 338 mN in the Z-axis, and 130 mN of shearing force in the X and Y axis across all egg models.

DISCUSSION

The developed magnetic force sensor demonstrates promising capabilities for integration into ophthalmic surgical forceps. The sensor's compact form factor enables integration with existing forceps designs.

Validation and Performance: The tape peeling trials successfully demonstrated consistent force measurements with repeatable performance across multiple trials. The sensor reliably captured three degrees of freedom, providing comprehensive force feedback during the controlled manipulation task. This repeatability validates the robustness of the calibration approach and the neural network mapping, establishing confidence in the sensor's real-time measurement capabilities during surgical manipulation. Biological validation through egg peeling experiments further demonstrated the sensor's capability to provide feedback during delicate tissue interaction.

Novelty and Clinical Significance: This represents the first implementation of direct grip force sensing at this scale integrated into the tip of micro-forceps. While FBG sensors have been integrated along instrument shafts to measure bending moments, they do not provide direct force sensing at the tissue contact point. Our force sensor directly measures grip force at the site of tissue manipulation, addressing a critical gap in robotic surgical feedback. This development enables future implementation of gripping haptic control for surgeons and is particularly valuable for retinal surgery, where delicate neurosensory tissue requires precise force regulation to avoid irreversible damage. Furthermore, real-time grip force measurements will enable closed-loop control strategies to reduce tissue stress through regulated gripping forces, potentially improving surgical outcomes and tissue preservation.

CONCLUSION

We have successfully designed and validated a magnetic force sensor for ophthalmic surgical robot forceps. The sensor provides real-time force feedback suitable for delicate surgical tasks, representing an important step toward enhanced force control in robotic ophthalmic surgery.

REFERENCES

[1] H. Mi, R. MacLaren, and J. Cehajic-Kapetanovic, "Robotising vitreoretinal surgeries," *Eye*, vol. 39, pp. 673–682, 2025.

[2] P. K. Gupta, P. S. Jensen, and E. de Juan Jr, "Surgical forces and tactile perception during retinal microsurgery," *Medical image computing and computer-assisted intervention–MICCAI'99*, pp. 1218–1225, 1999.

[3] N. Wang, X. Zhang, M. Li, A. Stilli, S. Bano, and D. Stoyanov, "FBG-based 3D micro-force sensor integrated into ophthalmic microsurgical robotic forceps," *Measurement*, vol. 271, p. 120909, 2026.

[4] I. Kuru, B. Gonenc, M. Balicki, J. Handa, P. Gehlbach, R. H. Taylor, and I. Iordachita, "Force sensing micro-forceps for robot assisted retinal surgery," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 1401–1404.

[5] M. Esfandiari, J. W. Kim, B. Zhao, G. Amirkhani, M. Hadi, P. Gehlbach, R. H. Taylor, and I. Iordachita, "Cooperative vs. Teleoperation Control of the Steady Hand Eye Robot with Adaptive Sclera Force Control: A Comparative Study," in *Proceedings - IEEE International Conference on Robotics and Automation*. IEEE, 2024, pp. 8209–8215.

A Concentric Tube Catheter, Actuated and Imaged via MRI

Peter Lloyd, Yael L. May, Jurgen E. Schneider, Pietro Valdastri

p.lloyd@leeds.ac.uk. STORM Lab, University of Leeds, UK. ePIC Lab, University of Leeds, UK

INTRODUCTION

Two major challenges associated with robotic catheterization are, firstly, the provision of controllable degrees of freedom (DoFs) and, secondly, accessing feedback on the shape and pose of the catheter [1]. Miniaturizable active steering can be achieved through magnetic actuation, and Magnetic Resonance Imaging (MRI) provides high definition, radiation-free 3D imaging that can be utilized for shape-sensing. Here, we propose a structurally adaptable Coaxial Sleeve Magnetic Actuator (CoSMA) [2], with deformation energy provided by the background field of the MRI scanner. Our approach combines the magnetic actuation principle of the easy axis of alignment [3] with the mechanical principles of concentric tube designs [4], [5]. This concept allows for a materially flexible, and therefore risk mitigated, multi-DoF catheter. We demonstrate the CoSMA (Fig.1), constructed of three coaxial components with respective outer diameters of 4 mm, 1.5 mm and 0.4 mm, in an aortic arch phantom navigation within the bore of a pre-clinical MRI scanner.

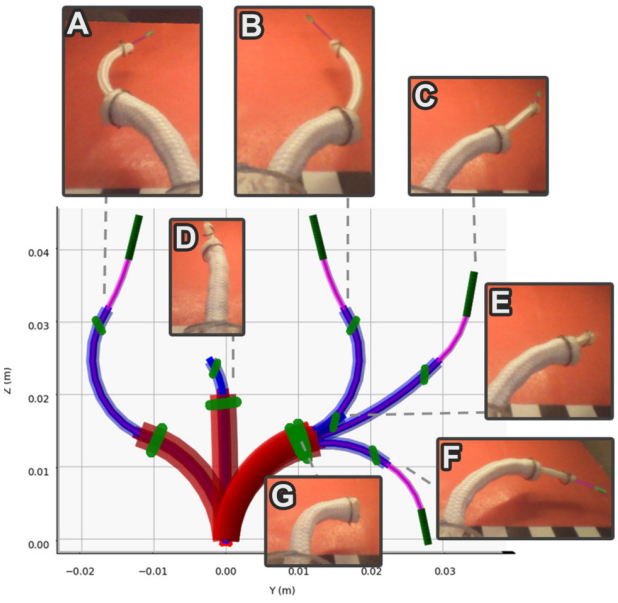


Fig. 1: Range of motion of the CoSMA in simulation and in the bore of the pre-clinical MRI scanner.

MATERIALS AND METHODS

Sleeves are listed from outer to inner - sleeve 1 has the largest diameter. For the inner sleeve (sleeve 3), 5 mm of 0.25 mm diameter iron wire (GF44272571 Iron Wire Reel

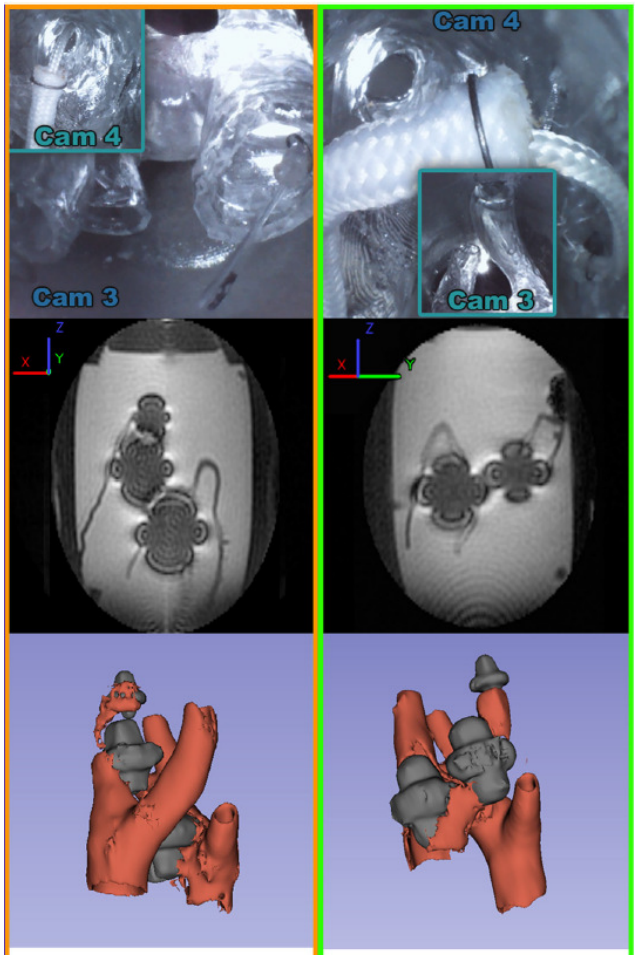


Fig. 2: Visual and MR Images of the aortic arch navigations - The right common carotid artery (RCA - orange) and the right subclavian artery (RSA - green). Optical camera images. 2D slice of the MR Image showing signal voids from ferrous components. Reconstruction of anatomy and CoSMA shape based on 3D MRI data.

- Merck KGaA, Darmstadt, Germany) was superglued in axial alignment to a 100 mm length of 0.4 mm diameter nylon wire. This was attached to an 800 mm long, 0.15 mm diameter nitinol rod. For the middle sleeve (sleeve 2), the same 0.25 mm diameter iron wire was wound once orthogonally onto a 2.1 mm diameter temporary rod. The rod ensures the windings will be circular (as opposed to elliptical) and thus exhibit no alignment bias when exposed to the background field. This ring was superglued at a 45° angle to a 1.5 mm diameter, 75 mm length, nylon braid. For

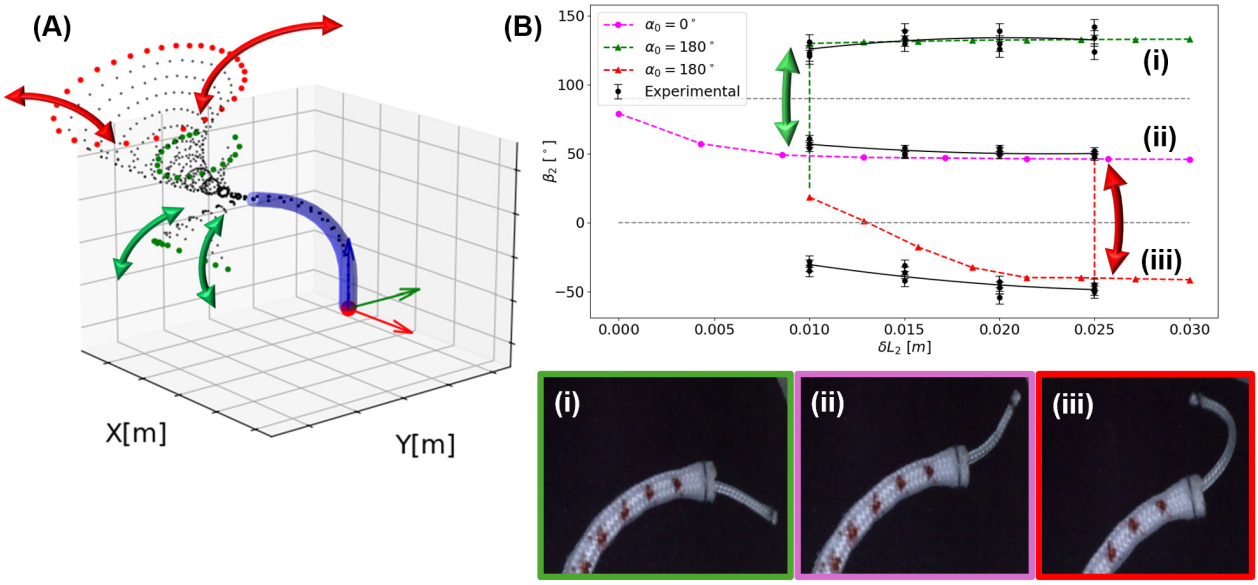


Fig. 3: (A) Range of motion plot based on simulation. (B) Numerical versus Experimental results. The initial pose at insertion follows the curve in magenta (ii). α_0 is rotated at $\delta L_2 = 0.01m$ to create the backward facing pose - the green curve (i). α_0 is rotated at $\delta L_2 = 0.025m$ to create the sigmoidal pose - the red curve (iii). Post rotation, sleeve 2 can be inserted and retracted without losing shape.

the outer sleeve (sleeve 1), a 50 mm long, 3 mm diameter braided nylon sleeve was threaded onto a temporary rod of 2.5 mm diameter. The same iron wire was wound once orthogonally around this sleeve and secured with superglue. The three sleeves are then threaded one onto the next in a concentric tube arrangement.

RESULTS

Fig. 2 shows visual and MR Images of two aortic arch navigations - the right common carotid artery (RCA) and the right subclavian artery (RSA). Optical camera images, 2D slice of the MR Image and the reconstruction of anatomy and CoSMA shape based on 3D MRI data are all shown (more details in [2]).

Fig. 3(A) shows a range of motion plot based on simulation [2]. Fig. 3(B) shows numerical versus experimental results - the initial pose at insertion follows the curve in magenta (ii). The base of sleeve 2 (α_0) is rotated at $\delta L_2 = 0.01m$ to create the backward facing pose - the green curve (i). The base of sleeve 2 (α_0) is rotated at $\delta L_2 = 0.025m$ to create the sigmoidal pose - the red curve (iii). Experimental results are shown in black. Post rotation, sleeve B can be inserted and retracted without losing shape.

CONCLUSIONS AND DISCUSSION

Our catheter design is capable of large multi-DoF deformations which are mechanically controllable from outside the MRI system. Evidently, CoSMA's body is significantly softer than would be possible without harnessing the magneto-static energy of the system. Furthermore, we have used anisotropic material composition (braided nylon) allowing maximization of the twisting to bending stiffness ratio (GJ/EI). This enhanced ratio reduces the tendency of the device to exhibit the snap-through instability [6]. We have demonstrated a potential clinical use case with manual navigation under visual feedback into the branches of the aortic arch. The principle benefit of magnetic actuation lies in the potential for miniaturization and the only barrier to significant improvement on our 4 mm maximum diameter design lies in fabrication. MR image distortion has been experimentally minimized and clearly image distortion/signal voids would reduce further for a more miniaturized system. The principle design feature of the CoSMA is that of retroflexion. The referential orientation of ring 1 is as close to 0° as possible (the symmetrical axis of the ring aligned with the background field) meaning sleeve 1 deforms to an angle asymptotically approaching 90°. Sleeve 2 can then rotate and/or translate to achieve a wide variety of shapes, for some of which the tip of sleeve 2 is deformed beyond 90°. This allows sleeve 3 to align either positively or negatively with the background field permitting the full range of motion demonstrated in Fig. 1.

REFERENCES

[1] T. da Veiga, J. H. Chandler, P. Lloyd, G. Pittiglio, N. J. Wilkinson, A. K. Hoshiar, R. A. Harris, and P. Valdastri, "Challenges of continuum robots in clinical context: a review," *Progress in Biomedical Engineering*, vol. 2, no. 3, p. 032003, aug 2020.

[2] P. Lloyd *et al.*, "A concentric tube catheter for endoluminal interventions, steered and imaged via magnetic resonance imaging." *Communications Engineering*, 2026.

[3] J. J. Abbott, O. Ergeneman, M. P. Kummer, A. M. Hirt, and B. J. Nelson, "Modeling magnetic torque and force for controlled manipulation of soft-magnetic bodies," *IEEE Transactions on Robotics*, vol. 23, no. 6, pp. 1247–1252, 2007.

[4] P. E. Dupont, J. Lock, B. Itkowitz, and E. Butler, "Design and control of concentric-tube robots," *IEEE Transactions on Robotics*, vol. 26, no. 2, pp. 209–225, 2009.

[5] R. J. Webster, A. M. Okamura, and N. J. Cowan, "Toward active cannulas: Miniature snake-like surgical robots," in *2006 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2006, pp. 2857–2863.

[6] D. C. Rucker and H. B. Gilbert, "Methods for improving stability of concentric tube steerable devices using asymmetric flexural rigidity," Apr. 23 2024, uS Patent 11,964,387.

Characterization of Robotic Needle Insertion for Automatic Biopsy Sampling of Dynamic Multilayer Polymeric Lung Phantoms

Maria Monserrat Diaz-Hernandez[1], Julian Ramirez-Nava[1], and Isaac Chairez[2]

Tecnologico de Monterrey, [1] School of Engineering and Sciences, and [2] Institute of Advanced Materials for sustainable Manufacturing

INTRODUCTION

Lung cancer remains one of the leading causes of cancer-related deaths worldwide, with approximately 1,817,469 deaths reported in 2022 [1], making early diagnosis essential for improving patient survival rates. Image-guided needle biopsy is commonly used for tissue diagnosis; however, during this procedure, specialists must manually insert the needle into the target to obtain tissue samples for pathological examination. This manual approach increases the risk of procedural errors and is further constrained by limitations in needle trajectory flexibility [2].

The implementation of robotic systems in clinical environments offers several advantages, including improved precision, accuracy, and reproducibility during tissue sampling; enhanced ergonomics; reduced operator fatigue; and decreased exposure of medical personnel to ionizing radiation [3]. Furthermore, robotic-assisted systems enable controlled and repeatable insertion trajectories and facilitate the integration of sensing technologies during needle-tissue interaction [4].

Robotic-assisted biopsy systems have emerged as a promising alternative to improve targeting accuracy and reduce human error during tissue sampling procedures. These systems enable controlled trajectories, repeatable insertion paths, and the possibility of integrating real-time sensing technologies [5]. Among these sensing modalities, force sensing has become increasingly important because insertion forces provide information regarding tissue interaction, penetration resistance, and possible material transitions.

For preliminary robotic validation, synthetic phantoms provide a controlled and reproducible environment in which material concentration, insertion depth, and layer interfaces are known a priori. Their purpose in this study is not to reproduce the complete biomechanical complexity of biological tissue, but to establish a controlled reference for evaluating robotic insertion and force acquisition before experiments in more complex biological models. Gelatin-based phantoms are particularly useful due to their low cost, ease of fabrication, and tunable mechanical properties [6].

The robotic positioning, orientation control, image-processing framework, and Non-Singular Terminal Sliding Mode (NTSM) controller used in this work were previously developed and validated in [7]. The present study does not re-evaluate these components; instead, it extends the platform through calibrated axial force sensing and quantitative analysis of robotic needle insertion in homogeneous and multilayer gelatin phantoms.

MATERIALS AND METHODS

The experimental platform consisted of a Cartesian XY positioning system and a myCobot 280 robotic manipulator, integrated with a custom needle holder and a force-sensing module, as shown in Fig. 1. The positioning and orientation procedures followed the previously validated robotic framework described in [7]. After reaching the target position, the robot adjusted the needle orientation and automatically executed the insertion and withdrawal trajectory.

In this study, the previously developed vision and control modules were used to provide reproducible needle positioning, while the experimental analysis focused specifically on insertion-force acquisition and its relationship with phantom composition and penetration depth.

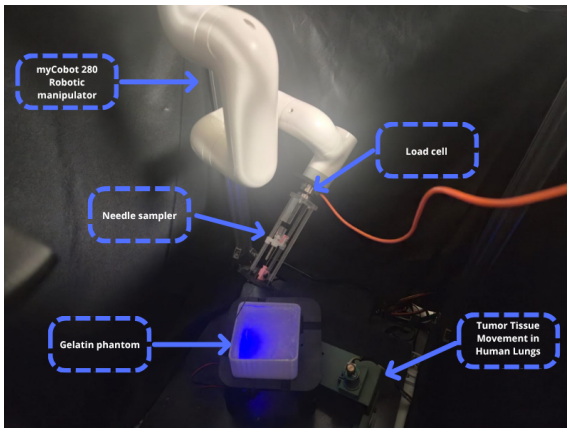


Fig. 1. Experimental robotic platform used for automated needle insertion and force acquisition. The setup includes the myCobot 280 robotic manipulator, biopsy needle holder, gelatin phantom, tissue target movement and load cell.

A. Force Sensor Calibration and Acquisition

Force measurements were obtained using a load cell connected to an HX711 amplifier and transmitted to MATLAB through serial communication. Before the insertion experiments, the sensing module was calibrated under axial loading over a range of 0–4.25 N. The calibration presented a linear response with $R^2 = 0.9993$, an RMSE of 0.0347 N, and an insertion-zero standard deviation of 0.0457 N. The measured signals were converted into axial force and synchronized with the robotic insertion trajectory. Additional no-contact trials were performed to characterize baseline disturbances before phantom insertion.

B. Gelatin Phantom Preparation

Two homogeneous gelatin phantoms were prepared at 10% and 20% (w/v) to characterize insertion behavior under uniform material conditions. A heterogeneous multilayer phantom was subsequently fabricated using sequential 10%, 20%, 10%, and 20% gelatin layers, each approximately 20 mm thick. This configuration produced known material interfaces at 20, 40, and 60 mm and enabled the measured force-profile changes to be compared with a geometric ground truth.

C. Insertion Protocol and Force Analysis

Before phantom insertion, robot translational motion was characterized under no-contact conditions using eight repetitions per command condition. SpeedCmd 6 was selected for the phantom experiments, corresponding to a measured velocity of 13.493 ± 0.480 mm/s.

Nine independent insertions were performed in each homogeneous phantom to a depth of 30 mm. For the multilayer phantom, nine insertion attempts were performed to 70 mm; two trials were excluded from quantitative analysis because the needle did not successfully penetrate the phantom, resulting in seven technically valid insertions.

The protocol consisted of moving to a safe position, adjusting needle orientation, approaching the phantom surface, performing the insertion trajectory, maintaining the predefined insertion depth, withdrawing the needle, and returning to the initial position.

Force profiles were analyzed as a function of insertion depth. For homogeneous phantoms, maximum force ($F_{\max}$), mean force over predefined depth intervals, and insertion work were calculated. For the multilayer phantom, local changes in the force profile were compared with the known interface positions at 20, 40, and 60 mm.

RESULTS

The 20% homogeneous phantom showed greater insertion resistance than the 10% phantom (Fig. 2(a)). Maximum force increased from 0.679 ± 0.134 N to 1.302 ± 0.381 N, while insertion work over 0–30 mm increased from 11.187± 3.545 mJ to 20.512 ± 9.558 mJ (Table I).

For the multilayer phantom, the seven valid insertions yielded $F_{\max}$ = 2.223 ± 0.103 N and 92.139 ± 5.519 mJ of work over 0–70 mm. As shown in Fig. 2(b), force-profile changes were detected at 20.286 ± 2.706, 42.214 ± 2.721, and 61.857 ± 2.868 mm for known interfaces at 20, 40, and 60 mm, respectively; localization errors are summarized in Table II.

TABLE I

HOMOGENEOUS INSERTION METRICS (MEAN ± SD).

Phantom	n	$F_{\max}$ **[N]**	F_{15-25} **[N]**	**Work**$_{0-30}$ **[mJ]**
10%	9	0.679 ± 0.134	0.526 ± 0.138	11.187 ± 3.545
20%	9	1.302 ± 0.381	1.023 ± 0.364	20.512 ± 9.558

TABLE II

MULTILAYER INTERFACE LOCALIZATION (n = 7).

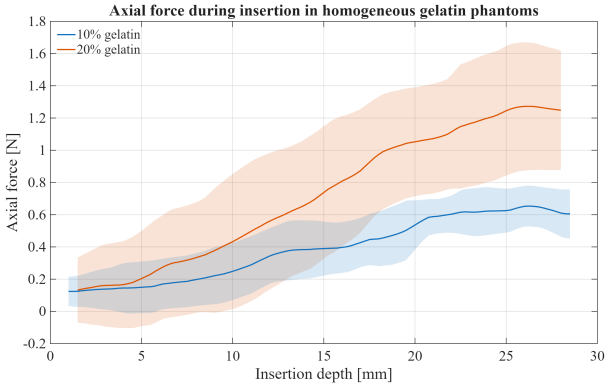
Transition	**Known [mm]**	**Detected [mm]**	**Abs. error [mm]**
10→20%	20	20.286 ± 2.706	2.143 ± 1.435
20→10%	40	42.214 ± 2.721	2.786 ± 2.018
10→20%	60	61.857 ± 2.868	3.000 ± 1.323

CONCLUSIONS AND DISCUSSION

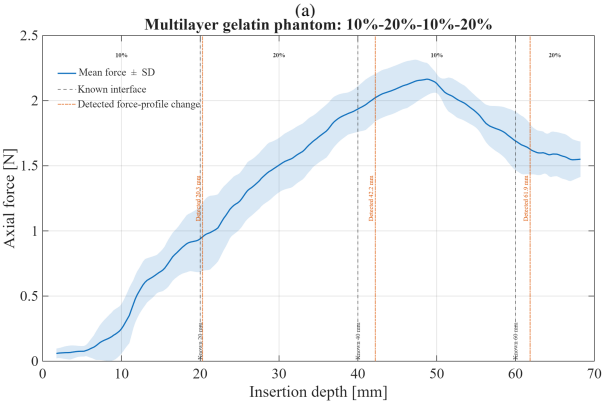
This study extends a previously developed robotic biopsy platform by incorporating calibrated axial force sensing and quantitative analysis of needle insertion in controlled gelatin phantoms. Under identical robotic conditions, the 20% homogeneous phantom exhibited greater insertion resistance than the 10% phantom, demonstrating that the sensing module is responsive to differences in phantom composition.

In the multilayer configuration, local changes in the force profile were spatially associated with known material interfaces, with mean absolute localization errors between 2.14 and 3.00 mm. However, the absolute force magnitude was not determined solely by local gelatin concentration, indicating that penetration depth, cumulative needle–phantom interaction, and global deformation also contribute to the measured signal. Therefore, the present results support force-based characterization of material transitions rather than direct tissue segmentation.

Future work will focus on improving contact-depth estimation and validating the force-based transition analysis in anatomically and mechanically more representative models.



(a)



(b)

Fig. 2. Force-depth profiles for (a) homogeneous 10% and 20% phantoms and (b) the multilayer phantom. Shading denotes ±SD; dashed lines indicate known and detected interfaces.

REFERENCES

[1] International Agency for Research on Cancer (IARC), "Cancer today: Pie chart data visualization," https://gco.iarc.who.int/today/en/dataviz, World Health Organization, 2022.

[2] C. de Margerie-Mellon, C. D. Bazelaire, and E. D. Kerviler, "Image-guided biopsy in primary lung cancer: Why, when and how," *Diagnostic and Interventional Imaging*, vol. 97, no. 10, pp. 965–972, 2016.

[3] S. M. Sajadi, S. M. Karbasi, H. Brun, J. Torresen, O. J. Elle, and K. Mathiassen, "Towards autonomous robotic biopsy—design, modeling and control of a robot for needle insertion of a commercial full core biopsy instrument," *Frontiers in Robotics and AI*, vol. 9, 2022.

[4] Y. Zhou, K. Thiruvalluvan, L. Krzeminski, W. H. Moore, Z. Xu, and Z. Liang, "Ct-guided robotic needle biopsy of lung nodules with respiratory motion–experimental system and preliminary test," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 9, no. 3, pp. 317–330, 2013.

[5] F. J. Siepel, B. Maris, M. K. Welleweerd, V. Groenhuis, P. Fiorini, and S. Stramigioli, "Needle and biopsy robots: A review," *Current Robotics Reports*, vol. 2, pp. 73–84, 2021.

[6] A. Leibinger, A. E. Forte, Z. Tan, M. J. Oldfield, F. Beyrau, D. Dini, and F. Rodriguez y Baena, "Soft tissue phantoms for realistic needle insertion: A comparative study," *Annals of Biomedical Engineering*, vol. 44, no. 8, pp. 2442–2452, 2016. [Online]. Available: https://link.springer.com/article/10.1007/s10439-015-1523-0

[7] M. M. Diaz-Hernandez, G. Ramirez-Nava, and I. Chairez, "Target tissue identification based on image processing for regulating automatic robotic lung biopsy sampler: Onsite phantom validation," *Sensors*, vol. 26, no. 5, p. 1723, 2026.

TELEOPERATION OF A SIX-DOF PARALLEL MANIPULATOR FOR SPINAL DECOMPRESSION SURGERY — EVALUATION VIA EGGSHELL GRINDING EXPERIMENTS

Keng-Hua Yeh[1], Shih-Hsin Chen[2], Hsuan-Yu Chen[3]
[1]Point Robotics Medtech Inc., Taiwan (charlie.yeh@pointroboticsinc.com)
[2]Department of Power Mechanical Engineering, National Tsing Hua University, Taiwan(ernest.chen82@gmail.com)
[3]Department of Orthopedic Surgery, College of Medicine, National Taiwan University and National Taiwan University Hospital, Taiwan (hychen83@ntu.edu.tw)

INTRODUCTION

Spinal decompression surgery requires precise bone removal while avoiding damage to surrounding soft tissues. This work proposes a teleoperated robotic system based on a 6-DOF parallel manipulator with a hybrid force–position control strategy for safe and controlled bone removal. Bone surface geometry is reconstructed using radial basis function (RBF) approximation from scanned or depth data, enabling accurate interaction with non-flat surfaces. The system is validated through teleoperated eggshell grinding experiments, demonstrating stable low-force interaction and complete material removal without damaging the membrane, indicating its potential for safe robot-assisted decompression surgery.

MATERIALS AND METHODS

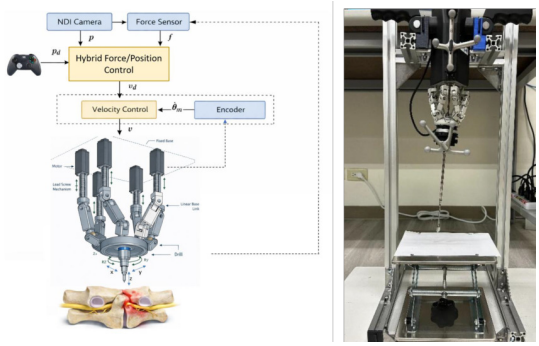


Fig. 1. System Architecture

The system is implemented using a six-degree-of-freedom (6-DOF) parallel manipulator equipped with a burr-type grinding end-effector. During operation, the operator teleoperates tool orientation and tangential motion along the surface, while penetration depth is automatically regulated through force control in the surface-normal direction. As illustrated in Fig. 1, the system adopts a closed-loop architecture integrating sensing, control, and actuation. Position feedback is obtained from an NDI optical tracking system, and interaction forces are measured by a force sensor mounted at the end-effector. These measurements, together with user commands, are processed by a hybrid force/position controller to generate desired velocity commands, which are executed by a low-level velocity control loop with encoder feedback to drive the parallel manipulator. The manipulator, kinematically similar to a Stewart platform, consists of six independently actuated limbs, each driven linearly by a rotary motor through a lead screw mechanism. The grinding tool, mounted on the moving platform, serves as the end-effector and enables precise regulation of both pose and contact force.

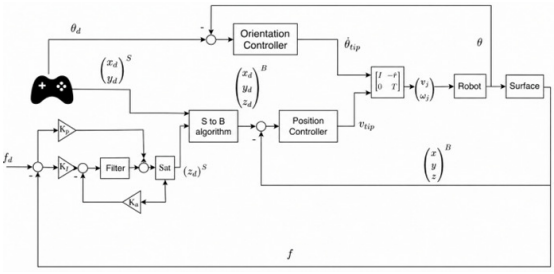


Fig. 2. Hybrid Force-Position Control

The hybrid force–position control scheme is illustrated in Fig. 2. Position control governs tool orientation and tangential motion along the surface, following teleoperation inputs, while force control regulates motion in the surface-normal direction using an outer–inner loop structure. The inner loop controls normal motion, and the outer loop employs a PI controller with anti-windup to generate a constrained position command based on a predefined maximum penetration depth. This constrained command is integrated into the overall motion control via coordinate transformation and position control. By decoupling tangential motion from normal force regulation, the system achieves stable surface tracking while maintaining a low and controlled contact force. The penetration constraint further prevents excessive tool intrusion, enabling safe, layer-by-layer material removal.

To obtain the bone surface topology required for robot grinding, a surface modeling approach based on RBF approximation is employed. Point cloud data acquired from manual scanning or a depth sensor are used to reconstruct the surface geometry, enabling estimation of surface normals and coordinate transformations for force control. As shown in Fig. 3, the eggshell experiments

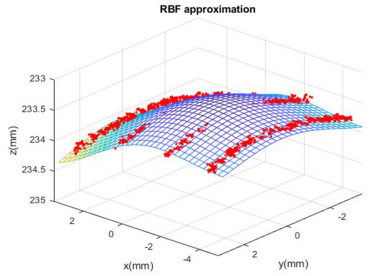


Fig. 3. RBF approximation using point cloud data from a depth sensor

achieve a reconstruction RMSE of 0.0943 mm, demonstrating sufficient accuracy for control.

RESULTS

The proposed method is validated through teleoperated eggshell grinding experiments that emulate bone removal in spinal decompression. The system maintains stable contact with a mean force of 0.5 N and a force-tracking RMSE of 0.142 N, while the depth-constrained control effectively prevents excessive penetration after shell breakthrough. To facilitate monitoring, a user interface provides real-time visualization of depth and trajectory; as shown in Fig. 4, the grinding area is discretized into a grid with each cell color-coded by penetration depth, enabling the user to identify processed regions, assess the depth distribution, and adjust the grinding trajectory accordingly. As shown in Fig. 5, the eggshell grinding experiment demonstrates successful material removal while preserving the membrane beneath.

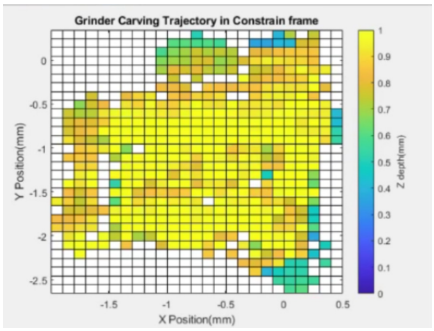


Fig. 4. User interface for real-time depth and trajectory monitoring

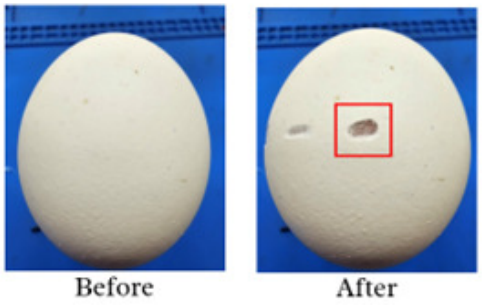


Fig. 5. Eggshell grinding results before and after material removal

Videos demonstrating the teleoperation and eggshell grinding experiments are available below:

Teleoperation video:
`https://youtu.be/l8R03s0SI2I`
Eggshell grinding video:
`https://youtu.be/uSOavtU1Dkk`

CONCLUSION AND DISCUSSION

This study presents a teleoperated robotic framework for safe material removal in decompression surgery. By integrating hybrid force–position control, RBF-based surface modeling, and depth-constrained regulation, the system achieves precise and stable interaction with complex surfaces. Experimental results demonstrate reliable low-force operation and safe layer-by-layer removal. The approach shows strong potential for robot-assisted spinal decompression and other applications involving delicate structures. Future work will focus on more complex anatomies and clinical validation.

REFERENCES

[1] Jin, H., Hu, Y., Li, F., and Zhang, J., 2011, "Safety Design and Control Algorithm for Robotic Spinal Surgical System," Proceedings of the First International Conference on Robot, Vision and Signal Processing (RVSP), IEEE, pp. 190–194.

[2] Li, M., Qi, X., Sun, Y., Li, B., Hu, Y., and Tian, W., 2023, "A Stability and Safety Control Method in Robot-Assisted Decompressive Laminectomy Considering Respiration and Deformation of Spine," ASME Journal of Automation Science and Engineering, 20(1), pp. 258–270. doi:10.1109/TASE.2022.3147270.

[3] Leite, A. C., Lizarralde, F., and Hsu, L., 2009, "Hybrid Adaptive Vision–Force Control for Robot Manipulators Interacting With Unknown Surfaces," The International Journal of Robotics Research, 28(7), pp. 911–926.

[4] Zhao, E., Marinho, M. M., and Harada, K., 2023, "Autonomous Robotic Drilling System for Mice Cranial Window Creation: An Evaluation With an Egg Model," Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 4592–4599.

[5] Kimura, K., Nakashima, S., Nakao, T., Inaba, M., and Okada, K., 2023, "Eggshell Cutting Task Strategy for a Robot That Handles Drill Based on Multiple Information Feedback of Force Sensing, Auditory Sensing, and Hand Position," Transactions of the Society of Instrument and Control Engineers, 59(1), pp. 37–47.

[6] Okuda, T., Yamashita, J., Fujita, M., Yoshioka, H., Tasaki, T., and Kato, A., 2014, "The Chicken Egg and Skull Model of Endoscopic Endonasal Transsphenoidal Surgery Improves Trainee Drilling Skills," Acta Neurochirurgica, 156, pp. 1403–1407.

[7] Andreoli, L., Simplício, H., and Morya, E., 2018, "Egg Model Training Protocol for Stereotaxic Neurosurgery and Microelectrode Implantation," World Neurosurgery, 111, pp. 243–250.

Compliance-Based Sensor Placement for Force Sensing on a Sensorized Prostate Phantom

Sizhe Tian, Yinoussa Adagolodjo, and Jeremie Dequidt

Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, University of Lille, Lille, France.

INTRODUCTION

Phantoms are widely used in medical training, such as for palpation-based examinations like the Digital Rectal Examination (DRE). Unlike traditional passive silicone phantoms, which provide only limited feedback [1], soft robotic phantoms equipped with sensing can quantitatively evaluate trainee performance [2], [3]. For this feedback to be meaningful, the phantom must estimate both where and how hard the trainee palpates, without restricting contact to pre-defined locations. The examination context imposes constraints: vision-based tracking may be occluded by the surrounding anatomy, and dense sensor arrays compromise the silicone's compliance and complicate fabrication. These constraints motivate a sparse sensing strategy, where sensor placement directly affects localization accuracy.

Existing methods typically maximize the information content of the configuration with respect to target modes or deformation patterns [4], [5], [6], [7]. We build on the QR-based sensor selection of Manohar *et al.* [4], but tailor it to soft palpation phantoms. The selection operates on a compliance matrix derived from FEM simulation, and a weighted greedy search prioritizes reconstructability within the clinically relevant DRE contact zone, while an exclusion zone keeps the direct contact interface itself sensor-free.

MATERIALS AND METHODS

The sensor set consists of two modalities: surface displacement markers whose placement we optimize via a compliance matrix, and three internal pneumatic chambers used as intrinsic pressure sensors. Surface markers directly measure local displacement; they are 2.5 mm passive retroreflective spheres tracked by an 8-camera Qualisys motion-capture system. The three-chamber design is inherited from our active phantom [8]. The chambers are unactuated, serving as passive sensors measured by MPXV7007DP differential pressure sensors.

External deformation induces volume changes in each internal chamber, resulting (via the isothermal Ideal Gas Law, $PV = \text{const}$) in measurable pressure variations. Chamber pressures thus provide a global, volume-derived signal that cannot localize deformation on its own, while the surface markers are local but spatially sparse; the two modalities are complementary.

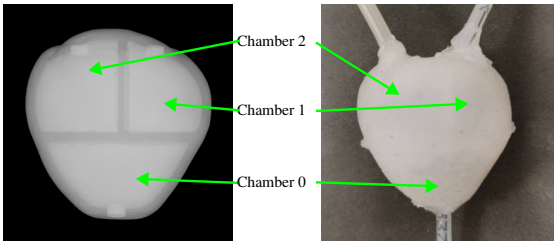


Fig. 1: Sensorized phantom: silicone body and finite-element method (FEM) model with three internal pneumatic chambers used as intrinsic pressure sensors. The phantom is fabricated by a lost-wax molding process using Ecoflex 00-10; fabrication details are available in our previous work [8].

Compliance Matrix

We generate a dataset by applying external forces in the three Cartesian directions at 1000 surface locations sampled by farthest-point-sampling, with FEM simulation in the Simulation Open Framework Architecture (SOFA) [9]. No internal actuation is applied, so the response captures passive compliance. Each surface node is treated as a candidate sensor location; for each sample we record node displacements and chamber pressures.

From the dataset, we construct the compliance matrix S, mapping force changes $\mathrm{d}F$ to measurement changes $\mathrm{d}Y$:

$$\mathrm{d}Y = S\,\mathrm{d}F.$$

S contains $3 + 3N$ measurement rows (3 chamber-pressure channels and 3 displacement components at each of $N = 4052$ nodes) and one column per sampled unit-force input. We scale pressure and displacement separately so each has a maximum sensitivity of 1, mitigating their unit-dependent magnitude difference (Pa vs. mm).

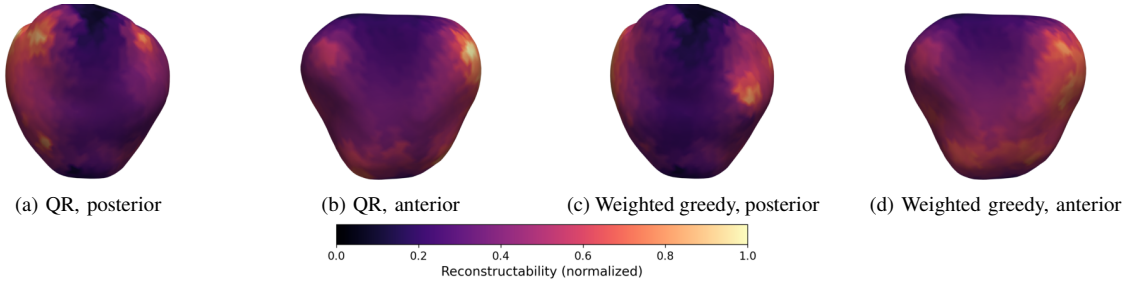
QR-Based Selection

As a baseline, pivoted-QR selection [4] is initialized with the three pressure sensors and greedily adds the surface node maximizing the residual norm $\|R_i^{(k)}\|_F$ after its three compliance rows are projected against the selected subspace $\mathcal{S}$, subject to the minimum distance $d_{\min} = 10\,\mathrm{mm}$.

However, global QR treats all surface regions equally; the resulting coverage is lowest on the posterior DRE region (Figures 2a and 2b), motivating a weighted variant that prioritizes a Region of Interest (ROI).

Weighted Greedy Search

For a candidate sensor set $\mathcal{S}$, let $S_{\mathcal{S}}^{(j)} \in \mathbb{R}^{|\mathcal{S}|\times 3}$ denote the local sub-compliance matrix at surface node j: its rows are



(a) QR, posterior (b) QR, anterior (c) Weighted greedy, posterior (d) Weighted greedy, anterior

Fig. 2: Reconstructability field with 10 markers selected by (a, b) QR and (c, d) weighted-greedy strategies, shown from posterior and anterior sides. Both fields are normalized by the maximum $\sigma_{\min}$ over surface nodes across both placements.

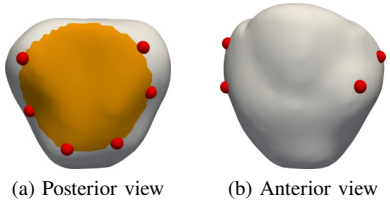
the measurement channels of $\mathcal{S}$, and its three columns are the unit-force responses at node j in the X, Y, Z directions. We define the per-node Reconstructability Score $\hat{s}_j(\mathcal{S}) = \sigma_{\min}\left(S_{\mathcal{S}}^{(j)}\right)$, the minimum singular value of $S_{\mathcal{S}}^{(j)}$ (the E-optimality criterion [10] applied per node); a higher score means forces applied at node j are more observable from $\mathcal{S}$. We define the ROI as the posterior surface region within 15 mm of the expected DRE contact center. Nodes inside the ROI receive weight $w_j = w_{\text{ROI}} = 5$; all others receive $w_j = 1$. To keep the direct contact interface unobstructed, we forbid sensor placement inside the ROI. Initialized with the three internal pressure sensors, the algorithm iteratively appends the candidate that maximizes the weighted-mean reconstructability over the mesh, subject to a minimum-distance constraint:

$$c^* = \underset{c \in \text{Cand}}{\operatorname{argmax}} \frac{\sum_j w_j \, \hat{s}_j(\mathcal{S} \cup \{c\})}{\sum_j w_j}$$
$$\text{s.t.} \quad \min_{s \in \mathcal{S}} \operatorname{dist}(c, s) \geq d_{\min}.$$

where Cand is the set of unselected surface nodes. The algorithm selects $K = 10$ surface markers which, with the 3 internal pressure sensors, form the 13-channel sensor set used throughout the remaining pipeline.

CONCLUSIONS AND DISCUSSION

With 10 surface markers and 3 pressure sensors, QR placement yields high reconstructability laterally but low values in the posterior DRE region (Figures 2a and 2b). Weighted-greedy placement prioritizes the ROI in its selection score (Figures 2c and 2d), concentrating observability there and increasing the mean ROI $\sigma_{\min}$ by 22.5% over QR.



(a) Posterior view (b) Anterior view

Fig. 3: Sensor configuration on the phantom: the 15 mm-radius ROI is orange, and the 10 selected surface markers are red.

The selected configuration (Fig. 3) reflects both mechanisms: the weighted score pulls markers toward the ROI, while the exclusion zone forces them to ring its boundary rather than sit on the contact interface; the anterior markers cover forces applied away from the DRE zone.

The current evaluation is only in simulation and treats the prostate in isolation. The model can be extended to surrounding anatomy by incorporating organ-contact constraints into the FEM compliance model and excluding or penalizing candidate marker locations likely to be occluded. Because sensor selection operates on the compliance matrix, it is not tied to optical tracking and can accommodate non-visual modalities, including the pneumatic pressure sensing used here and electromagnetic tracking, provided their measurement responses are included in the matrix. Future work will characterize sensor sensitivity, validate the placement on hardware and for downstream force localization, and extend the method to multiple or moving ROIs and other soft-body geometries.

REFERENCES

[1] G. J. Gerling, S. Rigsbee, R. M. Childress, and M. L. Martin, "The design and evaluation of a computerized and physical simulator for training clinical prostate exams," *IEEE Trans. Syst., Man, Cybern. A*, vol. 39, no. 2, pp. 388–403, 2009.

[2] S. Escaida Navarro, S. S. Dhaliwal, M. S. Lopez, S. Wilby, A. L. Palmer, W. Polak, R. Merzouki, and C. Duriez, "A bio-inspired active prostate phantom for adaptive interventions," *IEEE Trans. Med. Robot. Bionics*, vol. 4, no. 2, pp. 300–310, 2022.

[3] J. Hughes, P. Maiolino, T. Nanayakkara, and F. Iida, "Sensorized phantom for characterizing large area deformation of soft bodies for medical applications," in *Proc. IEEE Int. Conf. Soft Robot. (RoboSoft)*, 2020, pp. 278–284.

[4] K. Manohar, B. W. Brunton, J. N. Kutz, and S. L. Brunton, "Data-driven sparse sensor placement for reconstruction: Demonstrating the benefits of exploiting known patterns," *IEEE Control Syst. Mag.*, vol. 38, no. 3, pp. 63–86, 2018.

[5] J. Zhou, Z. Cai, P. Zhao, and B. Tang, "Efficient sensor placement optimization for shape deformation sensing of antenna structures with fiber bragg grating strain sensors," *Sensors*, vol. 18, no. 8, p. 2481, 2018.

[6] A. Spielberg, A. Amini, L. Chin, W. Matusik, and D. Rus, "Co-learning of task and sensor placement for soft robotics," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 1208–1215, 2021.

[7] D. Kim, S. Kang, and Y.-L. Park, "Optimal sensor placement for motion tracking of soft wearables using bayesian sampling," *Soft Robot.*, p. soro20240044, 2024.

[8] S. Tian, Y. Adagolodjo, and J. Dequidt, "Active prostate phantom with multiple chambers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2025, pp. 6904–6911.

[9] E. Coevoet, T. Morales Bieze, F. Largilliere, Z. Zhang, M. Thieffry, M. Sanz Lopez, B. Carrez, D. Marchal, O. Goury, J. Dequidt, and C. Duriez, "Software toolkit for modeling, simulation and control of soft robots," *Adv. Robot.*, vol. 31, no. 22, pp. 1208–1224, 2017.

[10] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.

Optimization-based compliance control for surgical continuum robots

Mattia Reina, Hassan Omran and Benoit Rosa
University of Strasbourg, CNRS, INSERM, ICube UMR 7357, Strasbourg, France

I. Introduction

Continuum robots, particularly concentric tube continuum robots (CTCRs), are flexible, slender structures capable of navigating constrained and tortuous anatomical environments. Their intrinsic compliance makes them suitable for minimally invasive procedures, allowing passive adaptation to contact and reducing the risk of tissue damage [1]. However, physiological motion, such as respiration or cardiac activity, may induce unintended interactions at arbitrary locations along the robot body. Sensorless techniques, including force estimation [2] and impedance control [3], have been proposed to regulate interaction forces. However, they either rely on model inversion and are therefore sensitive to modelling errors, or react only after contact occurs. Moreover, they are mainly formulated for tip interactions and do not readily extend to distributed contacts. Whole-body methods have also been investigated [4], but generally require accurate dynamic models, which are difficult to obtain for continuum robots because of nonlinear elasticity, friction, and configuration-dependent behaviour. In this work, we proactively reduce potential contact forces by exploiting the configuration-dependent compliance of the robot. Building on the generalized compliance formulation in [5], which enables efficient evaluation of compliance at arbitrary locations s along the robot body with respect to wrenches applied at s_0, we propose a quadratic programming (QP) controller that tracks a desired end-effector position while using kinematic redundancy to maximize directional compliance. The optimization direction encodes prior knowledge of likely contact conditions, such as anatomical constraints or physiological motion. The controller therefore selects configurations that reduce force transmission if contact occurs, without requiring measured interaction forces (Fig. 1). The proposed approach is evaluated under tip contact conditions.

II. Control Formulation

Optimization-based methods have already been adopted in continuum robotics, where redundancy and nonlinear mechanics make explicit control design difficult. Previous works have demonstrated that such formulations enable systematic handling of multiple tasks and constraints while ensuring stable behaviour [6]. In this work, we adopt a QP formulation. Directional compliance D is evaluated at an arc-length s along the robot body and with respect to external forces applied at a section s_0. It is computed from the generalized compliance matrix $\boldsymbol{C}_{s_0}(s)$ as

$$D(s) = \boldsymbol{v}^T \boldsymbol{C}_{s_0}(s) \boldsymbol{v} \tag{1}$$

where $\boldsymbol{v}$ is the direction of compliance maximization. This quantity measures how a force applied at s_0 produces a displacement at point s along direction $\boldsymbol{v}$. The proposed controller is inspired by [7] and extends it by explicitly incorporating compliance maximization as a control objective.

The problem is formulated as follows:

$$\min_{\dot{\boldsymbol{q}}, \boldsymbol{\delta}} \quad \frac{1}{2}\dot{\boldsymbol{q}}^T \boldsymbol{Q} \dot{\boldsymbol{q}} + \frac{1}{2}\boldsymbol{\delta}^T \boldsymbol{Q}_\delta \boldsymbol{\delta} - \alpha\tau(\nabla D)^T \dot{\boldsymbol{q}} \tag{2a}$$

$$\text{s.t.} \quad \boldsymbol{J}\dot{\boldsymbol{q}} + \boldsymbol{\delta} = \dot{\boldsymbol{x}}_d \tag{2b}$$

$$\boldsymbol{A}\dot{\boldsymbol{q}} \leq \boldsymbol{b} \tag{2c}$$

$$\dot{\boldsymbol{q}}^- \leq \dot{\boldsymbol{q}} \leq \dot{\boldsymbol{q}}^+ \tag{2d}$$

$$\boldsymbol{\delta}^- \leq \boldsymbol{\delta} \leq \boldsymbol{\delta}^+ \tag{2e}$$

The optimization simultaneously enforces two tasks. The first is a primary motion task (2b), ensuring tracking of the desired

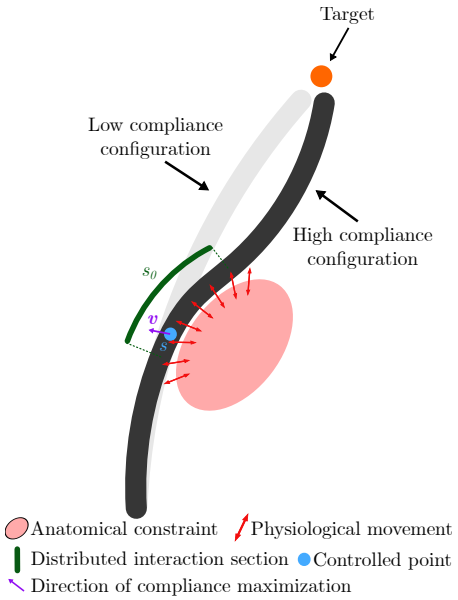


Fig. 1: Compliance-based configuration selection. Among configurations reaching the same target, the controller maximizes compliance along $\boldsymbol{v}$ to reduce contact forces. Here, s is an arc-length coordinate and s_0 is the wrench application section.

end-effector velocity $\dot{\boldsymbol{x}}_d$ through the robot Jacobian $\boldsymbol{J}$. This constraint is relaxed through a slack variable $\boldsymbol{\delta}$, which allows controlled deviations to keep the QP problem feasible. The secondary objective (2a) locally maximizes the directional compliance D with respect to the joint velocities, where $\dot{\boldsymbol{q}}$ denotes the control input and α weights its contribution. The parameter τ denotes the control loop time step and ∇D is the gradient of D w.r.t. $\boldsymbol{q}$, computed using finite differences. Constraints (2c-2e) implement standard actuator, velocity, and slack variable limits. By solving this problem at each control step, the robot configuration is selected to maximize compliance along a task-relevant direction while maintaining feasible motion.

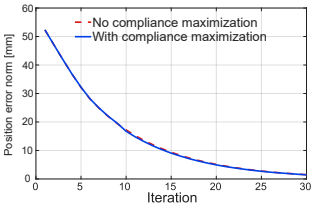
III. Simulation and Validation

We evaluate the proposed framework through simulation studies based on the physical parameters of a CTCR prototype available in our laboratory. A total of 500 trials were performed, each initialized from a different randomly sampled robot configuration. In each trial, the target was defined by displacing the initial end-effector position along a random unit direction by a distance randomly sampled between 0 and 10 cm. The parameter s_0 was set equal to the robot length, corresponding to interactions at the end effector, and $\boldsymbol{v}$ was a randomly sampled unit vector. Directional compliance was computed for both the proposed controller and a baseline position controller. The latter corresponds to the same optimization problem as in (2a-2e), with $\alpha = 0$, while the proposed method uses $\alpha = 1$. The compliance ratio is defined as the directional compliance at convergence for the proposed controller divided by that of the baseline. Across all trials, the proposed controller matched or outperformed the baseline in 100% of the cases, achieving a median compliance ratio of 1.50 and a mean ratio of 2.11. A representative trial is shown in Fig. 2.

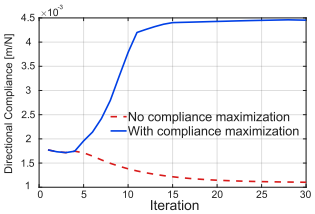
Directional compliance was experimentally evaluated on 30 representative robot configurations along each Cartesian direction X, Y and Z. For each configuration, a known external load was applied along the considered direction, and the resulting tip displacement was measured to estimate the directional compliance. The experimental compliance ratios were compared with their simulated counterparts, as shown in Fig. 3, confirming an overall increase in directional compliance along all three directions.

IV. Discussion

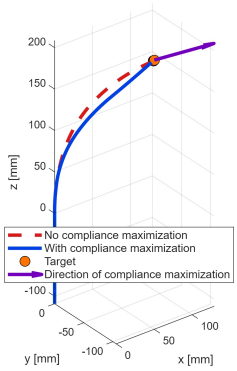
We proposed an optimization-based control strategy that maximizes directional compliance in order to minimize interaction forces in a likely direction of contact with the anatomy. Our method differs from prior impedance control approaches in that it does not explicitly rely on a dynamic model. Simulation results, supported by experimental validation, confirm that maximizing directional compliance at the control level effectively translates into increased physical compliance under external loading. Future work will extend the validation to distributed interactions along the robot body.



(a) Position error norm.



(b) Directional Compliance.



(c) Final robot configuration.

Fig. 2: Sample simulation trial over 30 iterations.

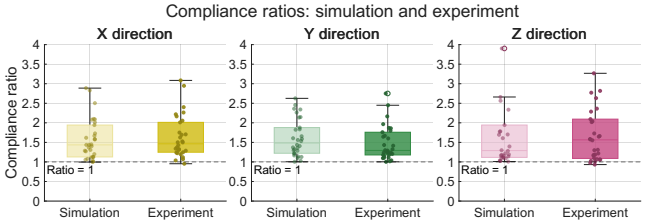


Fig. 3: Simulated and experimental compliance ratios over 30 configurations. The dashed line indicates the baseline.

Acknowledgments: This work of the Interdisciplinary Thematic Institute HealthTech, as part of the ITI 2021-2028 program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR-10-IDEX-0002) and SFRI (STRAT'US project, ANR-20-SFRI-0012) within the framework of France 2030.

References

[1] J. Burgner-Kahrs, D. C. Rucker, and H. Choset, "Continuum robots for medical applications: A survey," IEEE Transactions on Robotics, vol. 31, no. 6, pp. 1261–1280, 2015.

[2] M. Mahvash and P. E. Dupont, "Stiffness control of surgical continuum manipulators," IEEE Transactions on Robotics, vol. 27, no. 2, pp. 334–345, 2011.

[3] J. Duan, K. Zhang, K. Qian, J. Hao, Z. Zhang, and C. Shi, "An operating stiffness controller for the medical continuum robot based on impedance control," Cyborg and Bionic Systems, vol. 5, p. 0110, 2024.

[4] L. Toscano, V. Falkenhahn, A. Hildebrandt, F. Braghin, and O. Sawodny, "Configuration space impedance control for continuum manipulators," in Proc. 6th Int. Conf. Automation, Robotics and Applications (ICARA), 2015, pp. 597–602.

[5] G. Lods, B. Rosa, B. Bayle, and F. Nageotte, "Exact derivative propagation method to compute the generalized compliance matrix for continuum robots: Application to concentric tubes continuum robots," Mechanism and Machine Theory, vol. 200, p. 105696, 2024.

[6] T. Alsaka, P. Cinquin, and M. T. Chikhaoui, "Hierarchy control of dual-arm concentric tube continuum robots with different redundancy resolution techniques," in Proc. Int. Symp. Advances in Robot Kinematics, 2024, pp. 140–148.

[7] K. Dufour and W. Suleiman, "On maximizing manipulability index while solving a kinematics task," Journal of Intelligent & Robotic Systems, vol. 100, no. 1, pp. 3–13, 2020.

A Novel 2-DOF Continuous Hydraulic Motor for MR Intervention Robotics

Xinxin Li [1], Zion Tsz Ho Tse [1]

[1] *School of Engineering and Materials Science, Queen Mary University of London*

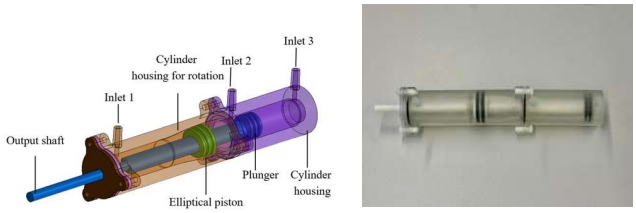
INTRODUCTION

X-ray, computed tomography (CT), and fluoroscopy have traditionally been used in diagnostic imaging and radiology for soft-tissue assessment. However, these modalities pose risks of ionising radiation, may yield suboptimal image quality, and can be challenging for physicians to interpret [1]. Magnetic resonance imaging (MRI) provides high-resolution, high-contrast visualisation of soft tissues. It enables real-time localisation of surgical tools and target anatomy, making it a powerful modality for minimally invasive interventions [2]. As MRI-guided procedures continue to expand in diagnostic and therapeutic applications, there is an increasing demand for robotic systems that can operate safely and precisely within the MR environment. However, developing high-fidelity MR-conditional actuators remains a significant challenge, as the limited availability of suitable actuation technologies constrains the performance and design of MRI-compatible surgical robots. In this article, we introduce a compact 2-DOF hydraulic motor capable of generating both rotational and linear motion. This motor offers a compact size, is MR-compatible, and is easy to assemble. Furthermore, its inherent 2-DOF capability facilitates integration into future MRI-guided robotic systems. The kinematic behaviour and transmission characteristics are analysed to support its control, and the motor's performance—including torque generation and hysteresis— is evaluated through a series of experiments.

MATERIALS AND METHODS

A. Motor Design

Fig. 1 (a) and (b) shows the detailed computer-aided design (CAD) model and a photograph, respectively. The motor has a cylindrical structure measuring 25 mm in diameter and 140 mm in length. The chamber used for elliptical piston rotation is formed by cutting an elliptical plane along a spiral line. For the rotational motion, to provide the rotational force, the channel through which the elliptical piston passes are also elliptical. As the elliptical piston rotates along a helical barrel, the tube and the plunger undergo synchronized rotational motion. When the tube undergoes linear motion, the elliptical-shaped tube slides through the elliptical piston, decoupling linear and rotational motion. The motor offers a theoretical maximum linear stroke of 33.25 mm and a rotational range of up to 101.7°.



(a) (b)

Fig. 1. Detailed 3-D drawing of the hydraulic motor. (b) Prototype of the motor.

To ensure MRI compatibility, key components—including the cylinder housing, tube, lid, and piston—were fabricated from three-dimensional (3D) printing resin, while screws were made of plastic. To allow for clear visualisation of piston movement and ensure the complete removal of bubbles from the chamber, thereby maintaining precise control, the cylinder housing was produced using Clear Resin V4. All parts were printed on a Form 3+ printer using stereolithography (SLA) technology, which facilitated the fabrication of robust and hermetically sealed components. For MRI compatibility, EPDM O-rings were used and housed in specially designed grooves to minimise water leakage.

B. Kinematics Analysis

The inverse kinematics of the 2-DOF hydraulic actuator establishes the relationship between the desired output displacements—angular and linear—and the fluid volume inputs required at each inlet. This formulation provides the mathematical foundation for displacement-driven control of the actuator, particularly in systems where direct position feedback may be unavailable or where precise volume control is the primary mode of actuation. Fig. 2 depicts the lateral view of the motor with kinematics parameters. All the parameters of the presented motor are given in Table I.

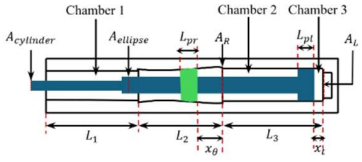


Fig. 2. Lateral view of the motor with annotated kinematic parameters.

TABLE I
DESCRIPTIONS OF PARAMETERS

Parameter	Description
$V_{3_initial}$	The initial water volume in chamber 1

V_1	The water volume in chamber 1
V_2	The water volume in chamber 2
V_3	The water volume in chamber 3
L_1	Length of the first cylinder chamber
L_2	Length of the helical chamber
L_3	Length of the second cylinder chamber
L_{exceed}	The length that the elliptical tube exceeds the first cylinder chamber when the piston for linear motion is in the rightmost position of the second cylinder chamber
L_{pr}	Height of the piston for rotational motion
L_{pl}	Height of the piston for linear motion
x_t	The translational displacement of the output shaft
θ	The rotational angle of the output shaft
L	The lead of the helical groove
A_L	Cross-sectional areas of the first- and second-cylinder chamber
A_R	Cross-sectional area for rotational motion
$A_{ellipse}$	Cross-sectional area of the elliptical rod
$A_{cylinder}$	Cross-sectional area of the cylindrical rod

The water volume in chamber 3 can be expressed as:

$$V_3 = V_{3_initial} + x_t A_L \tag{1}$$

The vertical component of the helical trajectory traced by the elliptical piston is given by:

$$x_\theta = \frac{L}{2\pi}\theta \tag{2}$$

Then the water volume in chamber 2 can be expressed as:

$$V_2 = (L_3 - x_t - L_{pl})(A_L - A_{ellipse}) + x_\theta(A_R - A_{ellipse}) \tag{3}$$

The water volume in chamber 1 can be expressed as:

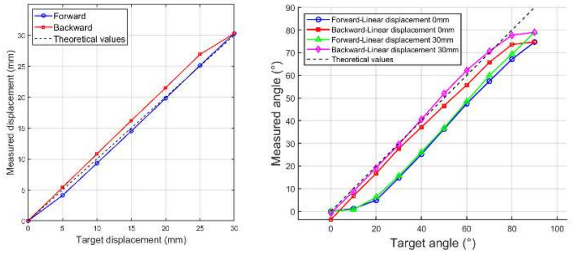
$$V_1 = (L_2 - x_\theta - L_{pr})(A_R - A_{ellipse}) + (x_t + L_{exceed})(A_L - A_{ellipse}) + (L_1 - x_t - L_{exceed})(A_L - A_{cylinder}) \tag{4}$$

To characterise the relationship between fluid injection and actuator motion, the partial derivatives of the chamber volumes V_1, V_2, V_3 concerning the translational displacement x_t and the helical displacement component x_θ were derived as following:

$$\begin{cases} \frac{\partial V_1}{\partial x_t} = A_{cylinder} - A_{ellipse} \\ \frac{\partial V_1}{\partial x_\theta} = \frac{L}{2\pi}(A_{ellipse} - A_R) \end{cases} \quad \begin{cases} \frac{\partial V_2}{\partial x_t} = A_{ellipse} - A_L \\ \frac{\partial V_2}{\partial x_\theta} = \frac{L}{2\pi}(A_R - A_{ellipse}) \end{cases} \quad \begin{cases} \frac{\partial V_3}{\partial x_t} = A_L \\ \frac{\partial V_3}{\partial x_\theta} = 0 \end{cases} \tag{5}$$

RESULTS

The forward and backward linear motion results are shown in Fig. 3(a). The maximum error was 2.2 mm with an average of 1.4 mm across all target points. The forward and backward rotational motion results are shown in Fig. 3(b). The rotational hysteresis at a linear displacement of 0 mm is a maximum of 21.9° and an average of 14.7°. When the linear displacement was increased to 30 mm, the maximum rotational hysteresis decreased to 15.1°, with an average of 9.8°. The results indicate that the longer the rod, the more pronounced the linear piston's influence, resulting in greater elastic torsion.



(a) (b)

Fig. 3. (a) Linear morion results. (b) Rotation results when the linear displacements are at 0 mm and 30 mm.

CONCLUSION AND DISCUSSION

To enhance the accuracy of the hydraulic motor, several design and control improvements can be implemented. These include increasing the stiffness of the connecting tubing, employing linear stages with motors capable of higher output forces, utilizing low-friction O-rings, and improving the surface finish of the cylinder housing's inner walls to reduce mechanical resistance. Furthermore, developing a motor force model and applying compensation-based control strategies can effectively mitigate hysteresis effects. A dynamic lag model combined with compensation control can also address the response delay observed in the current system. In addition, since the driving method follows an antagonistic actuation principle, the positioning accuracy can be further improved by increasing the pressure in both driving chambers, analogous to enhancing the preload force in antagonistic cable-driven systems. At present, the 3D-printed components are fabricated from resin, which exhibits relatively high friction with the sealing rings. For future product development, conventional syringe materials and industrial manufacturing methods may be adopted to reduce friction and improve overall performance.

REFERENCES

[1] M. U. Farooq and S. Y. Ko, 'A Decade of MRI Compatible Robots: Systematic Review', *IEEE Trans. Robot.*, vol. 39, no. 2, pp. 862–884, Apr. 2023, doi: 10.1109/TRO.2022.3212626.

[2] 'MRI Robots for Needle-Based Interventions: Systems and Technology | Annals of Biomedical Engineering'. Accessed: Dec. 11, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10439-018-2075-x

Elastic Loop Transfer Task: Bimanual Manipulation of Elastic Objects in Robot-Assisted Surgery

Shani Arusi[1], Uri Netz[1,2], and Ilana Nisky[1]

[1]*Department of Biomedical Engineering, Ben-Gurion University of the Negev*
[2]Department of Surgery A, Soroka University Medical Center

***Abstract*— Robotic systems have transformed many aspects of minimally invasive surgery, but the limited force feedback in many state-of-the-art systems limits the surgeon's ability to estimate interaction forces with soft tissue. Our previous work demonstrated systematic perceptual asymmetries in bimanual stiffness perception: elastic objects manipulated with the non-dominant hand are perceived as stiffer than those manipulated with the dominant hand. These findings encourage tailoring bilateral teleoperation controllers to human perception. The first step towards this goal is the design of bimanual tasks that require stiffness estimation and control, and that can, in the future, be performed with force feedback. In this work, we present one of them, the RAMIS Elastic Loop Transfer Task, in which users stretch an elastic loop using both instruments and place it between two pegs, mimicking surgical interactions with deformable tissues under tension. Ten participants performed the task using the da Vinci Research Kit without force feedback. We analyzed the resulting kinematics to assess bimanual coordination and task performance, providing a foundation of kinematic understanding for studying sensorimotor integration in RAMIS. By establishing this kinematic baseline, our work lays the groundwork for systematically evaluating how force feedback, and its tailoring to known perceptual asymmetries, shape bimanual surgical performance, ultimately informing the design of more intuitive and effective teleoperation systems for RAMIS.**

INTRODUCTION

Robot-assisted minimally invasive surgery (RAMIS) improves dexterity and visualization but still has limited force feedback. Consequently, surgeons rely mainly on visual information to infer tissue properties such as stiffness, particularly during manipulation of elastic tissue. Because many surgical procedures require coordinated use of both hands, understanding bimanual interaction strategies during elastic manipulation is essential.

Most existing evaluation tasks focus on general dexterity and do not explicitly examine bimanual interaction with elastic elements or the functional roles of each hand. Previously, we found that elastic objects manipulated with the non-dominant hand are perceived as stiffer than those manipulated with the dominant hand [1]. These findings encourage tailoring bilateral teleoperation controllers to human perception. The first step towards this goal is the design of bimanual tasks that require stiffness estimation and control.

In this study, we present one such task: the RAMIS Elastic Loop Transfer Task, implemented using the da Vinci Research Kit to evaluate bimanual interaction with elastic objects in a clinically relevant scenario. Users simultaneously stretched an elastic loop using both instruments and placed it between two pegs without force feedback. We analyzed the resulting kinematics and video data to assess bimanual coordination and task performance across trials and hand strategies, providing a foundation of kinematic understanding for studying sensorimotor integration in RAMIS.

METHODS

A. Participants

Ten right-handed participants (2 male and 8 females, aged 25.8 ± 4 years old) participated in the experiment. Handedness was assessed using the Edinburgh Handedness Inventory (scores above 40 were classified as right-handed). This study was approved by the Human Subjects Research Committee of Ben-Gurion University of the Negev, and participants were naive to the robotic system.

B. Task Development and Design

Three bimanual experimental tasks were developed using the da Vinci Research Kit (dVRK) to study stiffness estimation in clinically relevant robotic-assisted minimally invasive surgery scenarios. The tasks were designed within the constraints of the dVRK workspace and implemented using a custom 3D-printed platform tilted by 22.5° to match the stereo endoscopic camera view.

In this study, we focused on the first task of a larger experimental protocol, the Elastic Loop Transfer Task, in which participants stretched a red elastic band between a central peg and either a right or left peg, forming a fully coupled bimanual interaction where both hands manipulate the same object and experience identical stiffness. Each participant performed 20 trials in this task, comprising two conditions: 10 trials involving the central-to-right peg configuration (shown in Fig. 1A) and 10 trials involving the central-to-left peg configuration. The elastic element consisted of a commercially available hair band ($3.6\,cm$ diameter) and custom 3D-printed pegs. The color contrast between the band and its environment enabled reliable image-based tracking, building on established peg transfer and motion assessment paradigms [2], [3].

C. Data Analysis

To evaluate participant strategy, we used image processing to track the elastic loop by fitting an ellipse to the element in each frame. The path of the ellipse center was extracted and used to compute displacement, accumulated path length, and

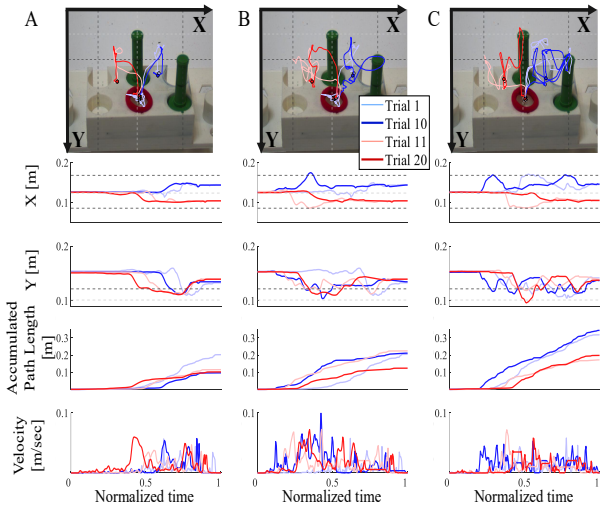


Fig. 1: (A-C) Performance of three of the participants. For each participant, the panels illustrate the spatial trajectory of the elastic loop center (top row), along with its X and Y coordinates, accumulated path length, and velocity as a function of normalized time. In all panels, light colors denote the first trial and dark colors denote the last trial within each section, while blue indicates trials from the section in which the peg is on the right side, and red indicates trials from the section in which the peg is on the left side. In the first 3 rows, the gray dashed lines indicate the centers of the green pegs.

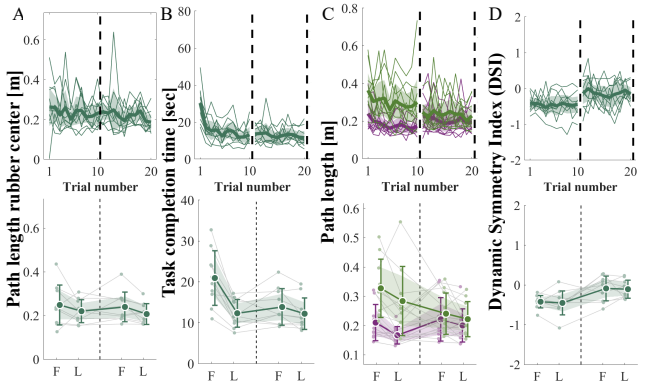


Fig. 2: Performance across trials: A) the elastic loop-center path length, B) task completion time, C) hands path length, and D) DSI between hands, all as a function of trial number. Upper panels show the learning curves; lower panels show the comparison between the first three (F) and last three (L) trials for each section.

its derivative profiles over time, characterizing the movement of the elastic element during the interaction.

In addition, we evaluated participant performance and learning using kinematic parameters derived from the dVRK instruments across trials. Task completion time was defined as the duration from the onset of movement until the participant indicated task completion. To quantify the difference between hands, we calculated the right- and left-hand path lengths (PL) over the task duration and computed a Dynamic Symmetry Index (DSI) based on the difference between them (Eq. 1, [4]). Positive DSI values indicated a
bias toward the right hand, whereas negative values indicated a bias toward the left hand. Larger absolute DSI values reflect greater asymmetry between hands.

$$DSI = \frac{\mathrm{PL}_{Right} - \mathrm{PL}_{Left}}{0.5 \cdot \left(\mathrm{PL}_{Right} + \mathrm{PL}_{Left}\right)} \quad (1)$$

RESULTS AND DISCUSSION

Participants revealed distinct coordination strategies for completing the task. The XY trajectories over time (Fig. 1 rows 2-3) highlight differences in coordination strategies between participants reflected in how the elastic center crossed the peg position (dashed gray lines). We observed strategy variability across participants: four participants stretched the elastic simultaneously with both hands (Fig. 1A), four participants (Fig. 1B) demonstrated a mixed strategy: early trials were dominated by first positioning the elastic center toward one peg, while later trials transitioned to a simultaneous stretching with both hands and two participant (Fig. 1C) consistently used a sequential strategy, placing the elastic on one peg before stretching it to the other peg.

Performance improved across trials. Elastic loop center trajectories became shorter (Fig. 1 row 4), and velocity increased in later trials (Fig. 1 row 5), indicating faster and more efficient control of the elastic object.

This overall reduction in path length across trials (Fig. 2A) indicates more efficient and confident manipulation of the elastic loop. Task-level performance metrics further support these observations. Task completion time (Fig. 2B) and path length of each hand (Fig. 2C) decreased across trials, indicating improved execution efficiency. Additionally, the DSI (Fig. 2D) remained negative across participants, indicating a greater path length for the left hand than the right, consistent with all participants being right-handed. These findings suggest that learning was accompanied by faster asymmetric execution despite the differences in learning strategies. Those results serve as a preliminary validation of the task, designed to later support force feedback evaluation.

REFERENCES

[1] S. Arusi, R. Leib, and I. Nisky, "Objects feel stiffer in the nondominant hand when comparing between hands," *Journal of Neurophysiology*, vol. 135, no. 4, pp. 1034–1048, Apr. 2026. [Online]. Available: https://journals.physiology.org/doi/10.1152/jn.00388.2025

[2] M. Hwang, J. Ichnowski, B. Thananjeyan, D. Seita, S. Paradis, D. Fer, T. Low, and K. Goldberg, "Automating surgical peg transfer: Calibration with deep learning can exceed speed, accuracy, and consistency of humans," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 2, pp. 909–922, 2023.

[3] F. Aghazadeh, B. Zheng, M. Tavakoli, and H. Rouhani, "Motion smoothness-based assessment of surgical expertise: The importance of selecting proper metrics," *Sensors*, vol. 23, no. 6, p. 3146, 2023.

[4] E. Galofaro, E. D'Antonio, N. Lotti, F. Patane, M. Casadio, and L. Masia, "Bimanual motor strategies and handedness role during human-exoskeleton haptic interaction," *arXiv preprint arXiv:2211.12408*, 2022.

Towards minimally invasive in situ bioprinting using a soft endoscopic robot

Ruben Martin-Rodriguez[1,2], Calvin Rüegg[1], Lorin Fasel[1], Nicolas Gerig[1], Valentina Basoli[2,3], Florian M. Thieringer[2,3], and Georg Rauter[1]

[1]*BIROMED-Lab, Department of Biomedical Engineering, University of Basel, Switzerland*
[2]*Swiss MAM Research Group, Department of Biomedical Engineering, University of Basel, Switzerland*
[3]*Clinic of Oral and Cranio-Maxillofacial Surgery and 3D Print Lab, University Hospital Basel, Switzerland*
Corresponding author: ruben.martinrodriguez@unibas.ch

INTRODUCTION

In situ bioprinting is a novel method leveraging the principles of tissue engineering to fabricate tissue and organ constructs directly at the defect site within the patient [1]. Compared with conventional methods, in situ bioprinting holds promise to improve implant integration, wound healing, and flexibility within the surgical setting. A major advantage is its potential for minimally invasive surgery, as the step of implantation is no longer required.

During minimally invasive bioprinting, end-effector motions are constrained by the insertion point, requiring additional degrees of freedom (DoF) inside the body to maintain a perpendicular orientation with the printing surface and ensure high print fidelity [2]. Several robotic systems and endoscopes for in situ bioprinting have been proposed to realize these additional DoF [3]. However, these systems are either voluminous, limiting their use to natural orifice insertion [4], or sharply bend the material transfer tube, impacting the material flow [5].

Here, we present a 1-DoF proof-of-concept endoscopic robot for minimally invasive bioprinting (Figure 1). The proposed design combines a continuum robot to minimize variations in material flow and series elastic actuation (SEA) to realize a soft robot with increased inherent safety. The proposed system was designed considering the requirements for bioprinting articular cartilage in the knee joint, a tissue with poor self-regeneration and often treated arthroscopically.

MATERIALS AND METHODS

For our endoscopic robot, we chose a planar, 1-DoF, tendon-driven continuum robot. The prototype design consisted of a 3D-printed monolithic TPU (TPU 95A HF, Bambu Lab, Shenzen, China) structure of 7 mm in diameter and 18 mm in length (Figure 1c), including a 0.5 mm PLA end plate (PLA Matte, Bambu Lab, Shenzen, China). A central channel was designed for the PTFE material feeding tube, with an outer diameter of 1.5 mm and an inner diameter of 0.84 mm (21G). The TPU structure featured a central backbone that supported the PTFE tube, constrained the length of the robot, and provided stability against out-of-plane motions. Fin-like structures on either side guided the tendons with a 3 mm lever arm for controlled bending.

The robot was actuated with an antagonistic SEA mechanism to increase safety and reduce tendon slack and backlash (Figure 1b). We used an impedance control for the end-effector position with a cascaded force controller for the individual force tendons [6]. To perform the forward and inverse kinematics, we used a constant-curvature model for its simplicity and compatibility with our robot design [7].

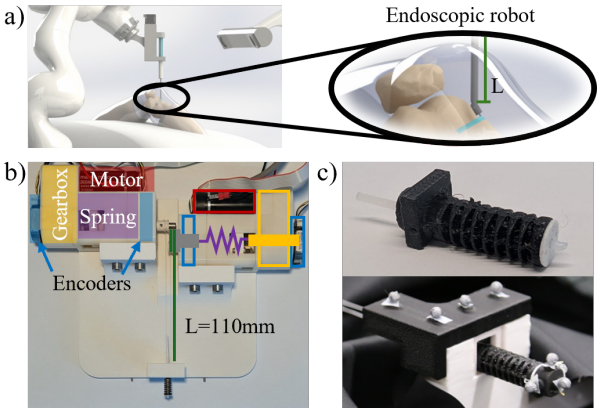


Fig. 1: Overview of endoscopic bioprinting using an endoscopic robot. a) Schematic representation of endoscopic knee cartilage bioprinting by combining a manipulator with an endoscopic robot. b) Antagonistic series elastic actuation of the robot with a schematic overlay of the actuation unit. c) Prototype of the soft endoscopic robot, and during evaluation using motion tracking. Scale bar: 0.5 mm.

We evaluated the static positioning accuracy of the robot by commanding end-effector angles θ_{des} in the range $[-50°, 50°]$ in 10° steps, starting from a straight configuration, $\theta_0 = 0°$, and from either end of the range of motion, $\theta_0 = \{-50°, 50°\}$. Each desired position was reached 10 times from each starting configuration. Robot motions were measured using a motion capture system (Miqus M5, Qualisys, Sweden). We calculated the error $\Delta\theta = \theta - \theta_{des}$ considering the end-effector angle computed from the tendon encoders and measured by the tracking system. We also calculated the deviations of the end-effector out of the plane of motion.

We also tested the printing capability of the robot. We selected a curved substrate following the theoretical end-effector position under a constant curvature assumption. This substrate was positioned at a distance of 0.4 mm from the end of the material transfer tube. A line of

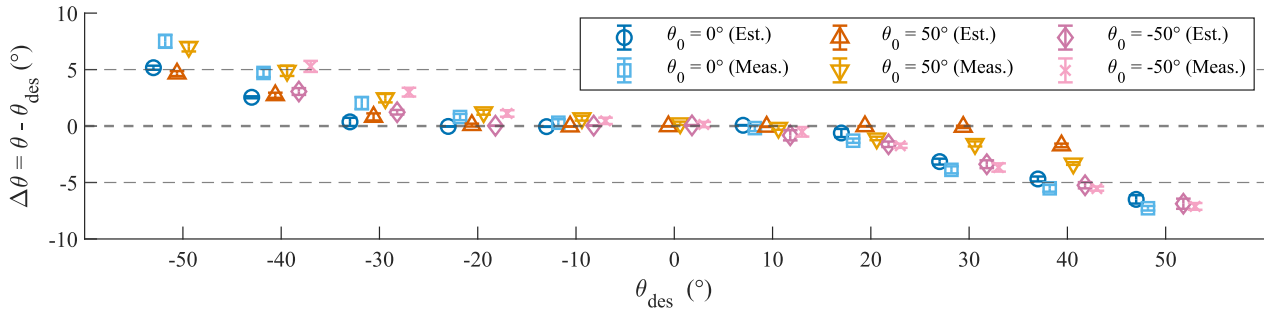


Fig. 2: Results of the static position accuracy evaluation. For each desired end-effector angle, θ_{des}, we computed the angular error, $\Delta\theta$, estimated from the encoders of the actuation units and measured by the tracking system (denoted *Est.* and *Meas.*, respectively). Each θ_{des} was reached from starting angles, $\theta_0 = [-50, 0, 50]$. $n = 10$ for all conditions.

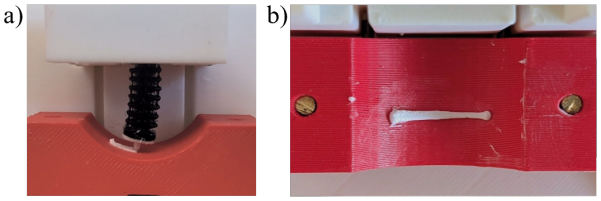


Fig. 3: Endoscopic 3D printing demonstration. a) Printing process. b) Resulting print onto the curved substrate.

commercially available cream as a viscoelastic gel model was printed at a speed of 3 mm s^{-1} (11 ° s^{-1}) from −40° to 40° via mechanical extrusion using a syringe pump (PSD/8, Hamilton Company, USA).

RESULTS

The positioning error across all conditions was 1.86 ± 0.72° and 2.83 ± 0.54° for the estimated and measured end-effector angles, respectively, with a maximum error of 8.94° when $\{\theta_{des} = -50°, \theta_0 = 0°\}$ (Figure 2). The difference between estimated and measured angle was 0.51 ± 0.58° across conditions, with a maximum discrepancy of 2.37° when $\{\theta_{des} = -50°, \theta_0 = 50°\}$. The deviations out of the plane of motion were 0.0512 ± 0.0581mm throughout the evaluation, with a maximum of 0.1124 mm. Movements were measured with the optical tracking system, providing a 0.0364 mm residual error in the region of interest.

The system was able to print a continuous line of material (Figure 3). Visual inspection showed a homogeneously printed line without perceivable deviations out of the desired plane. Excess material accumulated at the start and end of the print, due to a missing synchronization between robot motion and material flow. We observed a progressive decrease in line width due to a lateral misalignment between the robot and the substrate.

DISCUSSION AND CONCLUSION

Considering a nozzle diameter of 0.84 mm (21G) and the dimensions of our robot, an end-effector accuracy of 1.78° would be required to ensure continuous filament formation. To meet this requirement, we need to further improve our robot, particularly for angles larger than 30°. The out-of-plane motions of the robot throughout the evaluation remained mostly below 0.1 mm, which is sufficient for printing with nozzle diameters down to 0.25 mm.

The higher errors at end-effector angles over 30° suggest a need for further tuning of the controller. We observed good conformity between the estimated and measured angles, suggesting that the constant curvature assumption can be a valid model of our robot. Larger errors were observed when moving in the negative direction. We attributed this to higher tension on the antagonistic tendon, caused by higher friction in one of the SEA units due to prototype manufacturing inconsistencies

During the printing demonstration, we observed continuous and consistent material deposition. We considered that the larger curvature radius of our continuum soft robot helped to reduce the impact of tube bending on material flow, improving print consistency. Together with the SEA, our design can achieve a high safety level upon contact with sensitive tissue. When combined with an extracorporeal positioning system, it would enable printing into skeletal tissues, but also near and onto damaged or sensitive tissue, such as internal organs or the brain, broadening the scope of application of endoscopic 3D printing.

REFERENCES

[1] Y. Wang, R. Pereira, C. Peach, B. Huang, C. Vyas, and P. Bartolo, "Robotic in situ bioprinting for cartilage tissue engineering," *International Journal of Extreme Manufacturing*, vol. 5, no. 3, 2023.

[2] R. Martín-Rodríguez, G. Rauter, N. Sharma, V. Basoli, F. M. Thieringer, and N. Gerig, "A robotic platform for navigated intraoperative 3D printing of patient-specific implants," *at - Automatisierungstechnik*, vol. 73, no. 12, pp. 951–965, Dec. 2025.

[3] M. Maintz, Y. Tomooka, M. Eugster, N. Gerig, N. Sharma, F. M. Thieringer, and G. Rauter, "In situ minimally invasive 3D printing for bone and cartilage regeneration - a scoping review," *Current Directions in Biomedical Engineering*, vol. 10, no. 2, pp. 66–70, 2024.

[4] W. Zhao and T. Xu, "Preliminary engineering for in situ in vivo bioprinting: A novel micro bioprinting platform for in situ in vivo bioprinting at a gastric wound site," *Biofabrication*, vol. 12, no. 4, p. 045020, Aug. 2020.

[5] A. Simeunovic and D. J. Hoelzle, "Coupled Dynamics of Material Delivery and Robotic Manipulator Axes in Endoscopic Additive Manufacturing," in *2019 American Control Conference (ACC)*. Philadelphia, PA, USA: IEEE, Jul. 2019, pp. 687–692.

[6] L. Fasel, N. Gerig, P. C. Cattin, and G. Rauter, "Control Evaluation of Antagonistic Series Elastic Actuation for a Robotic Endoscope Joint," *Journal of Bionic Engineering*, vol. 19, no. 4, pp. 965–974, Jul. 2022.

[7] P. Rao, Q. Peyron, S. Lilge, and J. Burgner-Kahrs, "How to Model Tendon-Driven Continuum Robots and Benchmark Modelling Performance," *Frontiers in Robotics and AI*, vol. 7, p. 630245, Feb. 2021.

AR Visualization of Workspace Overlap for Collision Prevention in Robot-Assisted Surgery with Hugo™ RAS

Jonas Jakobi[1], Lukas Freiling[1], and Do Yeon Kim[1]

[1]*Fraunhofer Research Institution for Individualized Medical Technology and Engineering IMTE*

jonas.jakobi@imte.fraunhofer.de

INTRODUCTION

Robot-Assisted Surgery (RAS) has proven to be a viable approach for minimally invasive procedures with the potential of improving patient outcomes with faster recovery times [1]. Systems such as the Hugo™ RAS (Medtronic, USA) exemplify this advancement, enabling complex interventions with high precision while being ergonomic to use for the surgeon [2]. However, a recurring operational challenge is the potential external collision between robot arms during surgery. The instrument track of the robotic arms might collide, potentially damaging hardware as well as restricting the surgeon's ability to operate [3]. Preventing these collisions is difficult, as the arm setup must balance tool reach, camera angles, and spatial constraints imposed by trocar placement. Additionally, once the robot arms are set up, the requirements of surgery do not permit taking collision avoidance into account during operation, as many surgeons do not notice robot arms colliding until movement is noticeably restricted [4]. Therefore, it is necessary to preemptively avoid collisions by placing the robotic arms and trocars to minimize the likelihood of collisions in the first place, without compromising the freedom of motion of the surgical tools [5]. While AR has already shown promise in assisting the initial spatial setup and positioning of robotic components [6], currently, surgical teams lack intuitive feedback regarding the spatial reach of each arm once docked. To potentially mitigate collision risk, a system can be implemented for computing and displaying the range of motion (RoM) of the robotic arms, as well as the intersection of these ranges. This work showcases an implementation of such a system that can compute both the RoMs and intersections and visualize them in-situ using an Augmented Reality (AR) headset.

MATERIALS AND METHODS

A. Range of Motion

After the robot arms are set up, the operational movement of each robotic arm is constrained to rotations about the Remote Center of Motion (RCM) at the trocar entry point. We model the RoM volume $V_{arm,i}$ of the i-th robot arm as a cone originating from the RCM and oriented according to the robot's setup configuration. Similarly, the trocar's angular constraints are represented by a cone $V_{trocar,i}$, oriented along the body surface normal at the entry point.

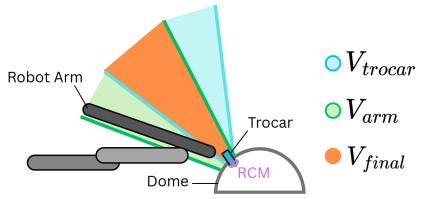


Fig. 1: Simplified view of the robots workspace V_{arm}, the trocars workspace V_{trocar} and the resulting operational range of motion V_{final}.

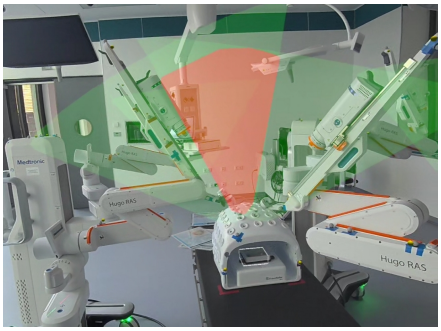


Fig. 2: View from an AR headset of the Hugo™ RAS surgical robot arms docked at a laparoscopic dome trainer. Green areas show the RoM volumes $V_{arm,i}$ of the robot arms and the red area the overlap volume $V_{collision}$.

As illustrated in Figure 1, the effective reachable workspace $V_{final,i}$ is determined by a Boolean intersection of these volumes:

$$V_{final,i} = V_{arm,i} \cap V_{trocar,i}.$$

The calculation is performed via constructive solid geometry (CSG), employing a Binary Space Partitioning (BSP)-Tree-based approach originally described by Thibault and Naylor [7] and implemented in the pb_CSG library. Finally, to identify the region of overlap, we compute the intersection area of the final RoM volumes between pairs of arms i and j. The total collision zone $V_{collision}$ of all robot arms is the union of all pairwise intersections:

$$V_{collision} = \bigcup_{i \neq j} (V_{final,i} \cap V_{final,j}).$$

Figure 2 shows such an intersection as the red area between the docked robot arms, as it is visualized on the AR headset.

B. Tracking and AR

To access the necessary pose information of the robot arms and the laparoscopic dome, we use an optical motion capture system (Qualisys AB, Sweden). All bodies of interest are equipped with retro-reflective passive markers, allowing for recording movements at 100 Hz using 18 infrared cameras. Marker poses were accessed using Qualisys Track Manager (QTM) and broadcast over the Robot Operating System 2 (ROS 2)[8], a middleware for robot communication and control.

The Meta Quest 3 (Meta Platforms, Inc., USA), a standalone mixed-reality headset with video passthrough, was used for AR visualization, running a custom application developed in the Unity game engine (Unity Technologies, USA). We use a serialized connection to a host computer running QTM and ROS 2, acting as a bridge to facilitate data transfer between the tracking system and the AR application. Upon starting the application, a user has to calibrate the coordinate frame of the headset with that of the tracking system by placing a virtual object at its corresponding real-world position, utilizing the laparoscopic dome as the reference. Through button presses on the Quest 3's controllers, real-time visualizations of $V_{arm,i}$ can be toggled on or off, and the computation and display of $V_{collision}$ can be initiated once the arms are docked.

RESULTS

The manual calibration between the AR visualizations and their corresponding real-world objects was evaluated by measuring the observable real-world gap. As the calibration is done at the laparoscopic dome trainer, three measurements were taken: a near measurement of offset error at the dome's marker, and two measurements of robot markers at different distances, to showcase the increasing inaccuracy when further away from the calibration point. The near measurement showed an alignment error of less than 1 mm, the first far measurement at 1 m distance an error of 2 mm, and the second far measurement at a distance of 1.47 m an error of 6.5 mm. The observed error gradient is characteristic of an angular displacement, leading to increased error at further distances.

Computational time was measured to evaluate the system's capabilities of real-time feedback. Rendering the current $V_{arm,i}$ of all robot arms can be done frame by frame without a measurable drop in framerate, as the cone's shape remains static and only moves along with the robot. The total time it takes to calculate $V_{collision}$ is relative to the number of arms involved in the calculation and can be seen in table I.

Number of Arms	Number of Overlaps	Total Time (ms)
1	0	178
2	1	349
3	3	651
4	6	865

TABLE I: Computational performance of the collision zone ($V_{collision}$) calculation based on the number of docked robot arms and their corresponding pairwise intersections.

CONCLUSIONS AND DISCUSSION

This work demonstrates the technical feasibility of an AR-based approach for real-time tracking and visualization of robotic workspace overlaps to mitigate intraoperative robot arm collisions. By overlaying three-dimensional volumes directly onto the physical surgical site, the system provides a more intuitive spatial context than 2D interfaces, potentially reducing the need for interpreting complex robotic constraints.

Although the alignment error was sufficient for this proof-of-concept, to overcome the inherent inaccuracy and variability of manual calibration, future iterations will integrate the AR headset directly into the tracking environment to ensure consistent spatial synchronization. The evaluation of computational load showcased the constraint of our current hardware and CSG-based operation. The final intersection volume $V_{collision}$ has to be computed on demand, which is currently sufficient for use at the laparoscopic training dome, as the assessment of collision zones happens after the arms are docked, at which point the positions are fixed, and continuous updates are not needed.

Transitioning to real patient anatomy would extend the possible trocar placement points into a continuous space. An efficient real-time calculation of $V_{collision}$ could then prove useful to continuously evaluate possible trocar positions and their resulting RoM overlap.

Future research will focus on a formal user study to determine if this intuitive spatial feedback can effectively reduce setup time and minimize the risk of collisions without compromising the surgical procedure.

REFERENCES

[1] S. Pasupuleti *et al.*, "The role of robotic systems in minimally invasive surgery: Benefits, risks, and future directions," *Int. J. Sci. Res. Eng. Manage*, vol. 5, pp. 1–6, 2021.

[2] F. Prata, A. Ragusa, C. Tempesta, A. Iannuzzi, F. Tedesco, L. Cacciatore, G. Raso, A. Civitella, P. Tuzzolo, P. Callè *et al.*, "State of the art in robotic surgery with hugo ras system: feasibility, safety and clinical applications," *Journal of personalized medicine*, vol. 13, no. 8, p. 1233, 2023.

[3] A. Gallioli, A. Uleri, J. M. Gaya, A. Territo, J. Aumatell, P. Verri, G. Basile, S. Fontanet, A. Tedde, P. Diana *et al.*, "Initial experience of robot-assisted partial nephrectomy with hugo™ ras system: implications for surgical setting," *World Journal of Urology*, vol. 41, no. 4, pp. 1085–1091, 2023.

[4] R. Randell, N. Alvarado, S. Honey, J. Greenhalgh, P. Gardner, A. Gill, D. Jayne, A. Kotze, A. Pearman, and D. Dowding, "Impact of robotic surgery on decision making: perspectives of surgical teams," in *AMIA Annual Symposium Proceedings*, vol. 2015, 2015, p. 1057.

[5] S. Gueli Alletti, V. Chiantera, G. Arcuri, A. Gioè, R. Oliva, G. Monterossi, F. Fanfani, A. Fagotti, and G. Scambia, "Introducing the new surgical robot hugo™ ras: system description and docking settings for gynecological surgery," *Frontiers in oncology*, vol. 12, p. 898060, 2022.

[6] D. Y. Kim, C. Adiyaman, M. Rolfs, G. Männel, and D. Kundrat, "Towards ar-based optimization of surgical robot training and setup," *Int. J. Sci. Res. Eng. Manage*, vol. 5, pp. 1–6, 2021.

[7] B. Naylor, J. Amanatides, and W. Thibault, "Merging bsp trees yields polyhedral set operations," *ACM Siggraph Computer Graphics*, vol. 24, no. 4, pp. 115–124, 1990.

[8] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, "Robot operating system 2: Design, architecture, and uses in the wild," *Science Robotics*, vol. 7, no. 66, p. eabm6074, 2022.

A Physics-Informed Neural Network for FBG-Based Force Vector and Location Identification along a Flexible Catheter

Xinran Liu[1], Mouloud Ourak[1], Johan Bennett[2], and Emmanuel Vander Poorten[1]

[1]*Department of Mechanical Engineering, KU Leuven University*
[2]*Department of Cardiovascular Medicine, University Hospitals Leuven*

INTRODUCTION

Chronic Total Occlusion Percutaneous Coronary Intervention (CTO-PCI) is among the most challenging procedures in interventional cardiology. Operators navigate guide guide catheters, microcatheters and dedicated guidewires — with working lengths close to one meter — through tortuous, calcified vessels, guided almost entirely by fluoroscopy. Tactile feedback at the proximal handle is strongly attenuated by shaft friction and vessel tortuosity, so the cardiologist effectively loses perception of contact forces along the distal shaft. Excessive, unperceived forces are a recognized cause of vessel dissection and perforation [1]. Due to the low instrument diameter (smaller than 3 mm outer diameter), embedding distributed force sensors to measure the interaction forces along the instrument shaft and at the tip is impractical. Instead, researchers have studied whether force could be estimated from the catheter shape [2], [3].

Multi-Core Fibers (MCFs), optical fibers with inscribed Fiber Bragg Gratings (FBGs), embedded along a flexible catheter provide distributed strain, and therefore curvature, measurements along its length [4]. However, recovering the cause of the observed shape — the magnitude, direction, and axial location of the contact force — is an ill-posed inverse problem. Where prior works recover force from shape using forward-mechanical (e.g. Cosserat-rod) models combined with iterative inverse solvers [5], such purely model-based approaches can be sensitive to parameter mismatch. Purely data-driven networks, on the other hand, typically require large labelled datasets and tend to generalise poorly outside their training distribution.

We therefore investigate a hybrid physics-informed neural network (PINN) [6] that learns the shape-to-force inverse mapping while embedding the Cosserat moment-balance equation directly in its training loss. The present study is a simulation-based feasibility study on a one-meter-long multi-core FBG catheter, restricted to planar deformation under a single perpendicular point force. Extensions to multiple, three-dimensional and distributed contact forces are discussed at the end of this abstract. In future work, real experimental data will be collected using FBG sensors for further investigation.

MATERIALS AND METHODS

Sensorized catheter model: The instrument is modelled as a cantilever Cosserat rod of length $L = 1000$ mm, diameter 3 mm, and bending stiffness $EI \approx 0.75$ N·m^2, which is within the range of a typical flexible stainless steel tube. It carries a multi-core FBG fiber with $N = 38$ gratings at a 26 mm pitch; the first grating sits at $s = 38$ mm (a 38 mm dead zone near the clamp), the last at $s = L$, for a total sensing length of 962 mm (37*26) and held horizontal along X axis.

Forward simulation: Training data are generated by solving the planar large-deflection Elastica equations for a perpendicular follower force of signed magnitude $F \in [\pm 0.3, \pm 1.3]$ N applied at an arc-length location $s_f \in [0.1\,L,\ L]$. The nonlinear boundary value problem (BVP) is solved by backward RK45 integration combined with a shooting method on the tip angle. The ground-truth curvature $\kappa(s)$ is then sampled at the center of the 38 FBG gratings, and an 2D shape is reconstructed from those samples via trapezoidal integration of $\theta(s) = \int \kappa\,(s)ds$, and then $(x(s), y(s)) = \int (\cos\theta(s), \sin\theta(s))\,ds$ — a constant-κ extrapolation is used in the proximal 0–38 mm dead zone. This approach mirrors the pipeline of a real FBG shape-sensing readout. A total of 3000 deflection scenarios (2400 / 300 / 300 train-val-test) are produced.

Network: A residual MLP (four hidden blocks of width 256 with LayerNorm + SiLU, ∼0.41 M parameters) maps the 38-dimensional curvature vector κ_{FBG} directly to two outputs: the signed force magnitude F (via tanh scaling) and the arc-length location s_f (via sigmoid scaling to $[0, L]$). Because κ is the primary sensor quantity, the network input depends only on localized FBG readings; the instrument shape in the XY-plane is a derived quantity used only downstream for visualization purposes. Shape mismatches due to integration errors will thus not impact the force and location prediction accuracy.

Physics-informed loss: The total loss is weighted by scaling factor β as $\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \beta\,\mathcal{L}_{\mathrm{phys}}$. The data loss $\mathcal{L}_{\mathrm{data}}$ is a smooth-L1 on normalized (F, s_f) with learned multitask uncertainty weighting to balance the two heterogeneous outputs. $\mathcal{L}_{\mathrm{phys}}$ enforces moment balance at every interior FBG node by comparing the internal bending moment carried by the rod, $M_{\mathrm{int}}(s) = EI\,\kappa(s)$, with the moment caused by the externally applied force, $M_{\mathrm{ext}}(s) = (\mathbf{r}_f - \mathbf{r}(s)) \times \mathbf{F}$ for $s < s_f$ and zero beyond, where $\mathbf{r}(s)$ is the rod position at arc length s and $\mathbf{r}_f = \mathbf{r}(s_f)$. The transition at s_f is blended by a smooth sigmoid to keep the loss differentiable in s_f. The external force is evaluated on the FBG-reconstructed shape, so inference requires only κ_{FBG}. β is held at zero for 50 warm-up epochs, linearly ramped over 50 epochs, and then fixed at 0.1. This is to prevent the network from steering the optimizer into a degenerate minimum by the physical loss. Training uses AdamW (LR 10^{-3}, weight decay 10^{-5})

with cosine-annealing warm restarts.

Force vector reconstruction: The full contact force vector is recovered at inference as $\mathbf{F} = F \cdot \hat{\mathbf{n}}(s_f)$, where $\hat{\mathbf{n}}$ is the unit vector perpendicular to the local tangent of the FBG-reconstructed shape at s_f.

RESULTS

Training converges in 129 s on a HP Zbook Fury 16 G11 with NVIDIA RTX 3500 Ada 12 GB graphics card. On the 300-sample held-out test set, the PINN recovers both force magnitude and application location with high accuracy:

- Error of signed force magnitude: Mean Absolute Error (MAE) = 15.6 mN (1.95% of the mean $|F|$ of 0.80 N); 50th / 90th / 95th percentile errors of 14.5 / 29.0 / 38.4 mN.
- Error of application location: MAE = 4.59 mm (0.46% of $L = 1$ m); 50th / 90th / 95th percentile errors of 3.91 / 8.90 / 11.30 mm.
- Error of reconstructed force components: F_x MAE = 5.8 mN, F_y MAE = 13.5 mN.

The larger F_y error is consistent with the dominant loading direction (perpendicular to the nominally straight catheter) and remains below 2% of the working force range.

Comparison with prior FBG-based force-from-shape methods: Table I positions the present single-force results against two previous researches [2], [5].The proposed PINN reduces both relative errors by roughly an order of magnitude (1.95% on $|F|$, 0.46% on s_f), and achieves this on a one-metre catheter — close to an order of magnitude longer than the instruments used in either reference. We stress, however, that the present numbers are obtained on noise-free, synthetically generated Cosserat curvatures, whereas the comparison works report results on physical hardware that includes FBG read-out noise, stiffness uncertainty and twist artefacts. Table I therefore provides an optimistic estimate of the performance of the physics-informed formulation.

TABLE I: Single-force estimation comparison. $|F|$ MAE percentages are normalised by the mean applied force; s_f MAE percentages are normalised by the instrument length L.

Method	$\|F\|$ MAE (mN/%)	s_f MAE (mm/%)
Qiao *et al.*	23.1 / 15.4	11.1 / 6.5
Lu *et al.*	27.8 / 12.4	3.8 / 3.2
This work (PINN)	**15.6 / 1.95**	**4.6 / 0.46**

As a first step towards multi-contact scenarios, we trained the same architecture on a three-simultaneous-point-force dataset. Accuracy drops relative to the single-force case, with an overall magnitude MAE of 0.21 N ($\approx$ 20% of the mean $|F|$) and an overall location MAE of 22.6 mm ($\approx$ 2.3% of L). The drop indicates that the current architecture and loss need further investigation on handling overlapping moment fields or the non-uniqueness of an ordered output representation when multiple forces act simultaneously.

CONCLUSIONS AND DISCUSSION

We have presented a physics-informed neural network that recovers a planar contact force vector and its application location along a one-metre multi-core FBG-sensed catheter, achieving sub-centimetre localisation and sub-2% magnitude error on the single-force test set under simulated FBG measurements. The study has two clear limitations: results are obtained on synthetic Cosserat-generated curvatures without added measurement noise; the three-force preliminary results highlight the limits of the current formulation on closely distributed forces.

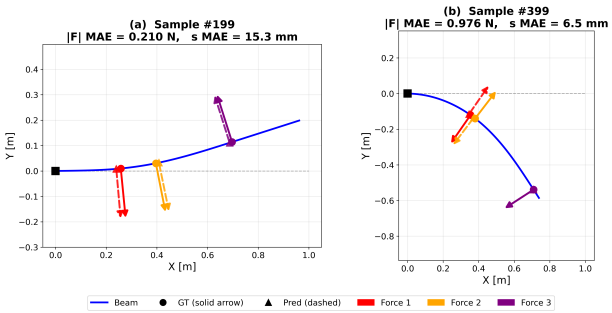


Fig. 1: Predicted vs. ground-truth three-force loading on two representative test samples.

Three directions are being pursued: (i) a permutation-invariant / energy-based reformulation of the multi-force loss to handle simultaneous contacts; (ii) extension of the Cosserat formulation from the planar Elastica to a fully three-dimensional rod, for arbitrary force directions and out-of-plane bending; and (iii) validation on data acquired with a real multi-core FBG-instrumented catheter on a benchtop silicone coronary phantom, addressing the synthetic-to-real gap and providing a first hardware demonstration along a clinical-length catheter.

ACKNOWLEDGMENT

This work was partially supported by RIVUS an internal KUL (C2E/25/024) and PROMPT. PROMPT received funding from the European Innovation Council (GA 101291374)

REFERENCES

[1] E. S. Brilakis *et al.*, "Guiding principles for chronic total occlusion percutaneous coronary intervention." *Circulation*, vol. 140, no. 5, pp. 420–433, 2019.

[2] Q. Qiao, G. Borghesan, J. De Schutter, and E. Vander Poorten, "Force from shape—estimating the location and magnitude of the external force on flexible instruments," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1826–1833, 2021.

[3] V. Aloi, K. T. Dang, E. J. Barth, and C. Rucker, "Estimating forces along continuum robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8877–8884, 2022.

[4] R. Roesthuis *et al.*, "Steering of multisegment continuum manipulators using rigid-link modelling and FBG-based shape sensing," *IEEE Transactions on Robotics*, vol. 32, no. 2, pp. 372–382, 2016.

[5] Y. Lu, B. Li, W. Chen, J. Yan, S. S. Cheng, J. Wang, J. Zhou, Q. Dou, and Y.-H. Liu, "Simultaneous estimation of shape and force along highly deformable surgical manipulators using sparse fbg measurement," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 9866–9872.

[6] M. Raissi *et al.*, "Physics-informed neural networks," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.

[7] D. C. Rucker *et al.*, "Statics and dynamics of continuum robots with general tendon routing and external loading," *IEEE Transactions on Robotics*, vol. 27, no. 6, pp. 1033–1044, 2011.

[8] O. Al-Ahmad *et al.*, "Fiber Bragg grating-based shape sensing for minimally invasive surgery," *IEEE Sensors Journal*, 2023.

Optimising Prostate Cancer Diagnosis with AI-Assisted Surface-Based Ultrasound-Derived Prostate-Specific Antigen Density

R.D. Bennett[1], T. Barrett[2], V. Gnanapragasam[2], Z. Tse[1]

[1]*School of Engineering and Materials Science, Queen Mary University of London, Mile End Road, London, UK, E1 4NS*

[2]*Department of Radiology/Urology, University of Cambridge School of Clinical Medicine, Cambridge, UK, CB2 0QQ*

z.tse@qmul.ac.uk

INTRODUCTION

Prostate cancer carries a substantial economic and patient burden [1]. However, nationwide screening is not currently recommended by the UK NHS due to limitations in the accuracy, cost, and tolerability of existing detection methods. Surface-based ultrasound-derived prostate-specific antigen density (SUS-derived PSAD) is a promising atraumatic alternative that agrees well with established imaging modalities [2], but currently relies on specialist expertise for prostate imaging and volume estimation. Here, we show that state-of-the-art deep learning frameworks can reduce this reliance while producing PSAD values that clinically agree with MRI-derived PSAD. Building on previous real-time SUS prostate tracking work [4], this study extends the approach to automated volume estimation, supporting AI-enabled SUS-derived PSAD as a potential triage tool to reduce unnecessary biopsies following PSA testing [3] in the diagnostic workup to prostate cancer.

MATERIALS AND METHODS

Datasets

A prospective clinical dataset of 19 patients was collected under ethical approval NRES 03/018. Each patient underwent transverse and sagittal transabdominal ultrasound (AUS) prostate imaging by an expert uro-radiologist, who also marked the central frame from each plane. MRI-derived prostate measurements and serum PSA values were available for all patients, allowing MRI-derived PSAD to be used as the gold standard.

PSAD Calculation

PSAD normalises serum PSA by prostate volume, with PSA obtained from blood testing and prostate volume estimated from imaging. In this study, volume was calculated using the ellipsoid approximation from three orthogonal dimensions (Eq. 1): right-left (RL), anterior-posterior (AP), and superior-inferior (SI). RL and AP were measured on transverse images, while SI was measured on sagittal images. Two thresholds were considered: the clinical threshold of 0.15 ng/mL^2 for MRI-derived PSAD and a conservative threshold of 0.12 ng/mL^2 for AUS-derived PSAD, chosen to account for the tendency of SUS-derived volumes to be overestimated [2].

$$PV = Rl \times AP \times SI \times C \qquad \text{Eq. 1}$$

Models Considered

One landmark identification model (IPV [5]), two object detection models (YOLO [6] and RetinaNet [7]), and one segmentation model (nnU-Net [8]) were evaluated for estimating the six prostate boundary endpoints used to derive the three dimensions in Eq. 1. Object detection-to-segmentation pipelines were also tested, where YOLO and RetinaNet outputs were used as localisation inputs to nnU-Net to improve segmentation performance, as shown in Figure 1. Models were evaluated using 5-fold cross-validation with patient-independent splits, kept consistent across object detection and segmentation models for fair comparison.

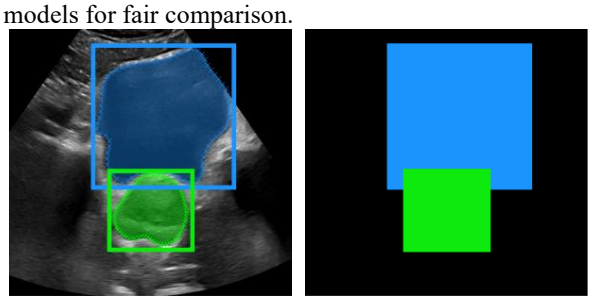
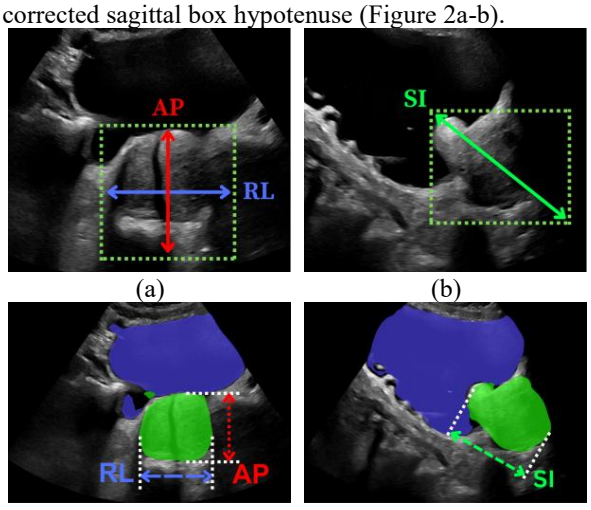


Figure 1: Detection boxes (left) to segmentation localisation inputs (right).

Model Output Conversions

Detection outputs were converted from bounding boxes into volume dimensions, with RL and AP taken from transverse box dimensions and SI estimated from the corrected sagittal box hypotenuse (Figure 2a-b).



(c) (d)

Figure 2: Conversion of detection outputs (top) and segmentation outputs (bottom) to ellipsoid dimensions.

For segmentation outputs, RL and AP were calculated from the maximum transverse mask extents, and SI from the upper-left-most to lower-right-most sagittal mask boundary points (Figure 2c-d).

RESULTS

Expert MRI-derived vs expert AUS-derived PSAD

Expert AUS-derived PSAD showed good agreement with expert MRI-derived PSAD (Figure 3). Using the clinical threshold alone, one classification disagreement was observed, indicated by the black star. Applying the conservative threshold to AUS-derived PSAD achieved 100% agreement with MRI-derived PSAD classification.

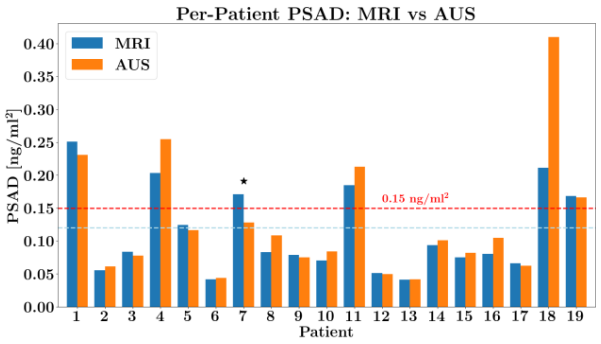


Figure 3: Expert MRI-derived vs expert AUS-derived PSAD at various threshold values for all 19 patients.

Model Metric Performance

Performance varied across endpoint localisation, dimension, and volume accuracy metrics. The RetinaNet+nnU-Net prostate-and-bladder pipeline achieved the best overall performance (blue), followed by the nnU-Net prostate-and-bladder model and IPV model (greens; Figure 4). Object detection models generally performed worse than segmentation-based approaches.



Figure 4: Model rankings based on average metric performance.

Model PSAD Performance

Segmentation-derived PSAD showed greater agreement with reference MRI-derived PSAD than object detection-derived PSAD (Figure 5). Object detection approaches tended toward false negatives, while segmentation approaches tended toward false positives. The only model achieving 100% agreement was the prostate-only RetinaNet+nnU-Net pipeline. The IPV model failed to calculate the volume for a single case, while all other models alwyas returned all dimensions, enabling volume calculations across all patients.

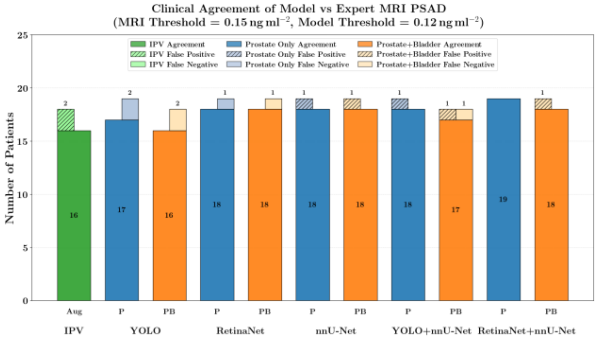
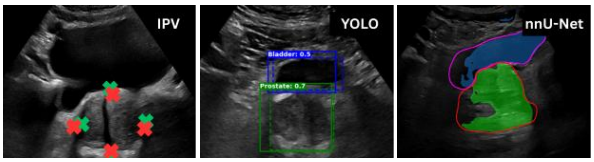


Figure 5: Clinical agreement between reference and model-derived PSAD.

Sample Outputs

Representative transverse outputs from the IPV, YOLO, and nnU-Net models are shown in Figure 6. IPV predictions were directly applicable to dimension calculation, while object detection and segmentation outputs first required conversion into prostate dimension estimates.



(a) (b) (c)

Figure 6: Sample transverse results from each model class before converting to dimensions.

CONCLUSION AND DISCUSSION

SUS-derived PSAD calculation can be largely automated using state-of-the-art deep learning frameworks. Segmentation models showed the strongest clinical agreement with MRI-derived PSAD, although faster object detection models still provide useful localisation inputs for downstream segmentation. False-negative classifications remain the key clinical concern, as they may lead to clinically significant prostate cancer being missed. These preliminary findings support the development of AI-enabled SUS-derived PSAD triage workflows to reduce unnecessary biopsies following PSA testing.

REFERENCES

1. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/mortality
2. https://doi.org/10.1371/journal.pone.0274014
3. https://www.baus.org.uk/patients/conditions/10/raised_psa/
4. https://doi.org/10.1038/s44172-025-00550-y
5. https://doi.org/10.3390/app12031390
6. https://github.com/ultralytics/ultralytics
7. https://doi.org/10.1109/TPAMI.2018.2858826
8. https://doi.org/10.1038/s41592-020-01008-z

Markerless Real-Time Target Tracking and Pose Estimation for Robot-Assisted Terahertz Imaging of Human Skin

Xiangyu Guo[1], Junlei Hu[2], Agrima Agarwal[1], Anubhav Dogra[1], Pietro Valdastri[2,*] and Emma Pickwell-MacPherson[1,*]

[1]Department of Physics, University of Warwick, Coventry, CV4 7AL, UK
[2]STORM Lab, University of Leeds, Leeds, UK
[*]Author to whom any correspondence should be addressed: E.MacPherson@warwick.ac.uk; p.valdastri@leeds.ac.uk

Introduction

Terahertz (THz) sensing is non-invasive and highly sensitive to water content in tissue, making it promising for assessing skin conditions and disease [1]. However, reliable measurements require normal probe placement and consistent contact pressure, which are difficult to achieve manually.

The PicoBot robot-assisted THz platform was developed to improve measurement repeatability [2]. PicoBot integrates a robotic manipulator, a Red-Green-Blue-Depth (RGB-D) camera, and force control to position the THz probe relative to the skin surface. In the previous workflow, the target region was localised using adhesive markers. This is inconvenient in clinical use, may be unsuitable near hair, wounds, or sensitive regions, and does not naturally support continuous tracking under patient motion.

In this work, we investigate whether PicoBot can be extended from marker-based localisation to markerless pose estimation for practical robot-assisted THz sensing. Building on previous markerless work focused on facial regions [3], this study extends the workflow to arbitrary accessible body regions. The key research question is whether a user-selected skin target can be localised, tracked, and converted into a stable robot-compatible probe pose without adhesive markers.

The main contribution is a general markerless localisation framework for robot-assisted THz sensing beyond facial scanning. The system combines interactive target selection, pre-contact target tracking, distance-aware local point-cloud processing, and robot frame pose estimation, enabling a more flexible markerless workflow for clinical measurements. Compared with related robot-assisted imaging systems using RGB-D sensing and point-cloud geometry for probe localisation [4], [5], this work focuses on a markerless THz workflow that removes adhesive markers while retaining practical probe-alignment accuracy.

MATERIALS AND METHODS

A. Markerless Target Localisation

As shown in Fig. 1, the pipeline starts from an operator-selected target point in the live RGB image. The selected pixel initialises CoTracker [6], which estimates the corresponding 2D target location in subsequent frames. This provides a continuously updated image-space target without adhesive markers or predefined landmarks.

B. Adaptive 3D Pose Estimation

For each tracked 2D target, synchronized RGB-D data are used to recover the corresponding 3D scan point. Given the camera intrinsics, the tracked pixel is back-projected into the camera frame to obtain the target position. Since the THz probe must approach the tissue at normal incidence, the probe orientation is estimated from the local surface geometry around this point rather than from the point position alone.

A local point-cloud neighbourhood is extracted using an adaptive CropBox centred at the reconstructed target. To maintain a consistent physical surface region under different camera-to-skin distances, the CropBox size is scaled with measured depth, $s = kd$, where d is the camera-to-target distance and k is an empirically selected scale factor. This distance-dependent region selection ensures that enough local points are retained for normal estimation while avoiding an overly large neighbourhood that would smooth over local curvature.

The cropped point cloud is filtered to reduce depth noise and outliers. Surface normals are then computed within the local neighbourhood, and quaternion averaging is applied to obtain a smooth probe orientation. The resulting pose is represented in the camera frame as a target position and a surface-normal-aligned orientation.

C. Robot Pose Transformation and Probe Positioning

The estimated target pose is transformed from the camera frame to the robot base frame using the calibrated eye-in-hand relationship. The final valid pose before probe descent is used as the reference for PicoBot probe positioning. During approach, the robot operates under Cartesian impedance control to maintain compliant and safe interaction with the patient.

The Robot Operating System (ROS) coordinates tracking, pose-estimation, and robot-motion nodes. Before the motion command is sent, a lightweight supervision layer checks whether the tracked target remains within the reliable RGB-D sensing range. If the target is occluded or the depth becomes invalid, the pose update is paused until a valid estimate is recovered.

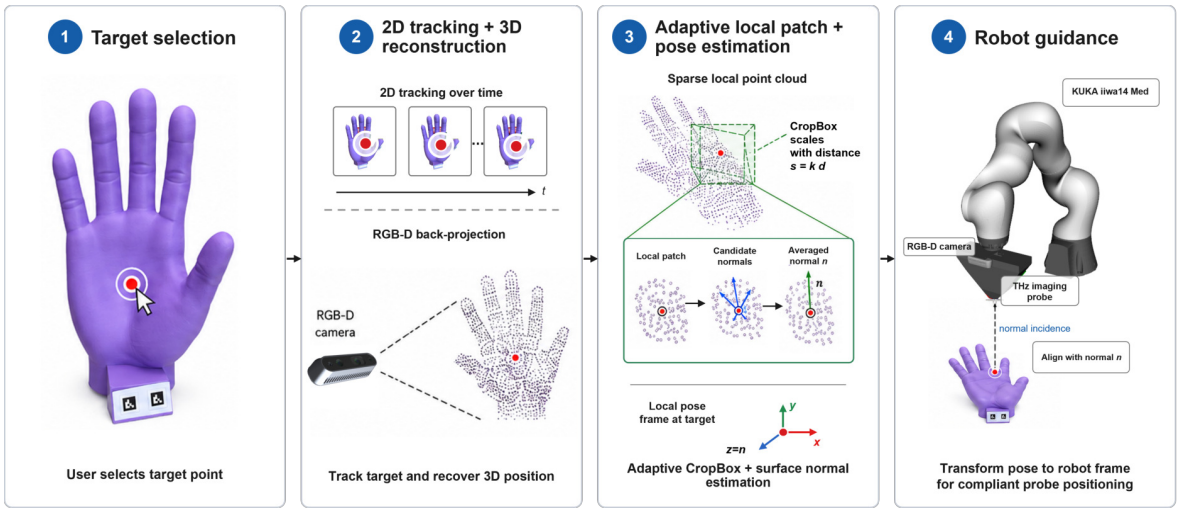


Fig. 1: Proposed markerless perception and robot-guidance workflow for PicoBot THz probe positioning.

TABLE I: Pose-estimation error against ArUco-registered CAD ground truth.

Point	e_x (mm)	e_y (mm)	e_z (mm)	$\|\mathbf{e}\|$ (mm)	Normal (°)
P1	+1.18	-1.30	-1.40	2.29	4.12
P2	+3.20	-1.55	-1.25	3.74	3.68
P3	+3.10	-1.90	-2.20	4.34	3.21
P4	+0.30	+3.20	-3.80	4.96	4.05
P5	+1.45	-0.20	-3.60	3.89	4.47
Mean	+1.85	-0.35	-2.45	3.84	3.91

Once the target pose is accepted, the probe is moved towards the skin and contact is regulated using a PID force controller combined with Cartesian impedance control. The desired contact force can be specified through the GUI before measurement and is limited to 0–5 N.

VALIDATION AND RESULTS

To evaluate pose-estimation accuracy for PicoBot guidance, we used a rigid PLA 3D-printed hand phantom. The phantom provides repeatable CAD-defined targets and curved regions for geometric validation, although it does not replicate soft-tissue deformation. Five points (P1–P5) were measured 60 times each at an RGB-D working distance of approximately 350 mm.

An ArUco marker (DICT_4X4_50, ID 0, marker length 18.8 mm) was used only to define the ground-truth reference frame. The marker pose transformed the CAD-defined target positions and surface normals into the camera frame, which were compared with the markerless estimates.

For this workflow, 5 mm and 5° were used as practical acceptance criteria, reflecting finite probe contact, force control, and the aim of reducing air gaps on curved skin. The mean position-error norm was 3.84 mm and the mean normal error was 3.91° (Table I), satisfying this criterion and supporting markerless guidance without adhesive markers.

CONCLUSIONS AND DISCUSSION

We developed a markerless framework for interactive target localisation and probe pose estimation in the PicoBot THz system. Validation on a rigid PLA hand phantom showed a mean position error of 3.84 mm and a mean normal error of 3.91°, satisfying the practical accuracy criterion for PicoBot probe guidance.

The complete PicoBot platform has also been used in hospital THz measurements involving 79 patients under TERABOTICS ethical approval (IRAS project ID: 270335), demonstrating stable operation during extended procedures. All research was performed in accordance with the relevant guidelines and regulations, including in accordance with the Declaration of Helsinki. Future work will validate the full workflow on soft tissue phantoms and quantify how residual pose error affects THz signal quality.

ACKNOWLEDGMENTS

This work was supported by the EPSRC Terabotics project, EP/V047914/1.

REFERENCES

[1] K. I. Zaytsev, N. V. Chernomyrdin, K. G. Kudrin, A. A. Gavdush, P. A. Nosov, S. O. Yurchenko, and I. V. Reshetov, "In vivo terahertz pulsed spectroscopy of dysplastic and non-dysplastic skin nevi," in *Journal of Physics: Conference Series*, vol. 735, no. 1. IOP Publishing, 2016, p. 012076.

[2] A. Dogra, D. Jones, A. I. Hernandez Serrano, S. Chakraborty, J. J. Young, B. G. Page, J. Hardwicke, P. Valdastri, and E. Pickwell-MacPherson, "Towards autonomous robotic thz-based in vivo skin sensing: the picobot," *Scientific Reports*, vol. 15, no. 1, p. 4568, 2025.

[3] S. Chakraborty, A. I. Hernandez-Serrano, A. Dogra, D. Jones, A. Agarwal, P. Valdastri, and E. Pickwell-MacPherson, "Markerless pose estimation and tracking on facial regions for in vivo robot-assisted terahertz sensing," *IEEE Transactions on Medical Robotics and Bionics*, 2026.

[4] X. Ma, Z. Zhang, and H. K. Zhang, "Autonomous scanning target localization for robotic lung ultrasound imaging," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 9467–9474.

[5] J. Tan, Y. Li, B. Li, Y. Leng, J. Peng, J. Wu, B. Luo, X. Chen, Y. Rong, and C. Fu, "Automatic generation of autonomous ultrasound scanning trajectory based on 3-d point cloud," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 4, pp. 976–990, 2022.

[6] N. Karaev, I. Rocco, B. Graham, N. Neverova, A. Vedaldi, and C. Rupprecht, "Cotracker: It is better to track together," in *European Conference on Computer Vision*. Springer, 2024, pp. 18–35.

Applying AI techniques to EMG sensor data to predict localized muscle fatigue in RAS training: Standing and sitting postural approach.

D. Caballero[1], A. Fernández-Montero[1], J.A. Sánchez-Margallo[1], F.M. Sánchez-Margallo[2]

[1]Robotics, bio-printing and Image-guided surgery Unit [2]Scientific Direction

Jesús Usón Minimally Invasive Surgery Center, Road N-521 Km. 41.8, ES-10071, Cáceres, Spain. {dcaballero, afernandez, jasanchez, msanchez}@ccmijesususon.com

INTRODUCTION

Robot-assisted surgery (RAS), offers several advantages for patients [1]. However, it presents significant physical and mental challenges for surgeons. Advancements in wearable technology facilitate the recording of ergonomic and physiological parameters of the surgeon during surgical training.

Artificial Intelligence (AI) has experienced exponential growth in recent years. AI algorithms are designed to uncover hidden knowledge within data, particularly throughout predictive models. The integration of wearable technology with AI holds substantial potential for developing novel methods to monitor, manage, and understand the relationship between surgeons' ergonomic risks and health risks [2].

The primary advantage of the current ergonomic risks monitoring method during surgical training lies in its predictive capabilities. By applying predictive models, evaluations are expedited, preventive, and objective compared to subjective methods that relay on human interpretation.

Consequently, this study endeavors to predict the physiological and ergonomic parameter of surgeons using wearable technology during conventional and RAS procedures. Moreover, this study aims to design and implement predictive models for the future prediction of musculoskeletal risks in surgeons, such as localized muscle fatigue, during surgical training activities.

MATERIALS AND METHODS

Experimental setup

The data for this study were collected during forty-two surgical training on RAS sessions, twenty-one of them of standing-RAS and the remaining ones of sitting-RAS. All surgical training sessions were performed randomly by seven surgeons experienced in laparoscopic surgery and low experience in RAS, two women and five men between the ages of 25 and 45, all of whom were right-handed.

The data were recorded using a portable, wireless EMG system that allowed the EMG signal to be recorded for subsequent calculation of localized muscle fatigue in eight muscle groups bilaterally.

On the other hand, current and immediately preceding situations were analyzed to implement predictive models of localized muscle fatigue. A general dataset (318,328 records) was generated for standing-RAS (159,164 records) and sitting-RAS (159,164 records). A preprocessing technique (scaling) was applied to these two datasets. After that, two new datasets were generated: 80% of the data for training purpose (127,331 records for each type of approach) and 20% of the data for testing (31,833 records for each type of approach). Records from all surgical training sessions were included in all datasets. Data from each muscle group bilaterally were added in each record.

Surgical training sessions

For RAS, the Versius® robotic platform (CMR urgical®, Cambridge, United Kingdom) was used for all surgical training sessions.

EMG wearable device

The data were recorded by using wireless EMG TRIGNO® Avanti system from DELSYS®. It was used to record the muscle activety by means of EMG signals. This system has up to sixteen sensors with a sampling rate of 1024 Hz. The EMG signal of following muscle groups was recorded bilaterally: upper trapezius, middle trapezius, erector spinae, brachioradialis, deltoids, brachii triceps, gastrocnemius medialis and vastus lateralis. The EMG sensors were located on each muscle according to the SENIAM guidelines [3]. The EMG values were presented as percentage of maximum voluntary contraction (%MVC). This was obtained separately for each muscle group just before each test by asking each surgeon to perform specific contractions against a fixed resistance in order to fix the performance of the baseline.

Localized muscle fatigue

For the analysis of localized muscle fatigue, the evolution of amplitude and median frequency of the muscle activity was analyzed for each task and surgical procedure. Localized muscle fatigue was quantified using the joint analysis of spectrum and amplitude (JASA) method [4].

Artificial Intelligence techniques

Seven different AI techniques were applied to the training dataset to generate predictive models based on different approaches: Convolutional Neural Network (CNN), Multilayer perceptron (MLP), Random forest (RF), Gradient reinforcement neural network (GBM), Multiple linear regression (MLR), Support vector machine (SVM) and U convolutional neural network (UNET). The Pytorch package v2.10, an AI library of the Python programming language v3.10.17. Finally, to validate the generated models, they were evaluated on the 10-fold cross-validation and test datasets. R^2

coefficient and Root Mean Square Error (RMSE) were employed to assess the predictive quality.

RESULTS

Training results

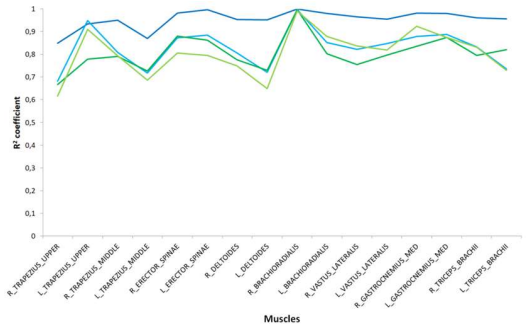


Figure. 1. Comparative plot of the best result on training dataset (MLR as AI technique) (NP-Standing RAS: dark blue; NP-Sitting RAS: light blue; Suturing-Standing RAS: dark green; Suturing-Sitting RAS: light green).

Figure 1 presents the predicted results for localized muscle fatigue regarding the muscle groups of the training dataset. In general, needle passing (NP) achieved better results than suturing; the standing-RAS achieved better results for NP and sitting-RAS for suturing. Right-hand muscles obtained better results than left-hand muscles. The highest results were achieved for MLR followed by GBM, RF, SVM, UNET, MLP and CNN.

Cross-validation results

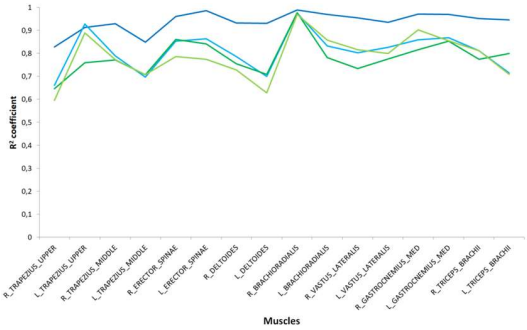


Figure. 2. Comparative plot of the best result on cross-validation results (MLR as AI technique) (NP-Standing RAS: dark blue; NP-Sitting RAS: light blue; Suturing-Standing RAS: dark green; Suturing-Sitting RAS: light green).

Figure 2 introduces the predicted results for localized muscle fatigue with regard to the muscle groups of the cross-validation results. In general, NP reached better results than suturing; the standing-RAS obtained better results for NP and sitting-RAS for suturing. Right-hand muscles obtained better results than left-hand muscles. The highest results were achieved for MLR followed by GBM, RF, SVM, UNET, MLP and CNN.

Test results

Figure 3 presents the predicted results for localized muscle fatigue on the muscle groups of the testing dataset.

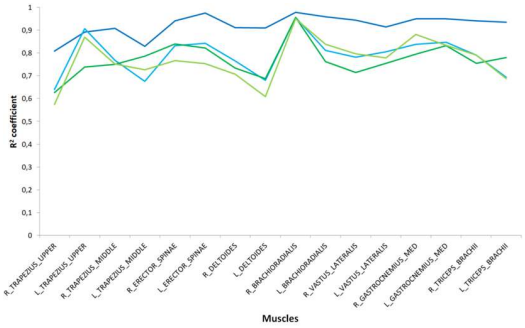


Figure. 3. Comparative plot of the best result on test results (MLR as AI technique) (NP-Standing RAS: dark blue; NP-Sitting RAS: light blue; Suturing-Standing RAS: dark green; Suturing-Sitting RAS: light green).

In general, NP obtained better results than suturing; the standing-RAS reached better results for NP and sitting-RAS for suturing. Right-hand muscles achieved better results than left-hand muscles. The highest results were achieved for MLR followed by GBM, RF, SVM, UNET, MLP and CNN. All muscles bilaterally showed statistically significant differences between sitting-RAS and standing-RAS ($p<0.001$) and between dominant and non-dominant side ($p<0.001$).

CONCLUSION

A novel predictive model based on MLR was developed to predict the localized muscle fatigue of the surgeon during RAS on surgical training tasks. This model achieved a satisfactory R^2 coefficient for all parameters. The models from machine learning and artificial neural network could learn erroneously in the approach used in current study. In further studies, additional AI algorithms must also be compared. Consequently, it represents a promising starting point for predicting surgeons' musculoskeletal risks.

Acknowledgments

This work was co-funded by the European Union through the Interreg VI-A Spain–Portugal (POCTEP) 2021–2027 Programme and the Junta de Extremadura (GR24210).

REFERENCES

[1] Caballero D, Sánchez-Margallo JA, *et al.*, "Applications of Artificial Intelligence in Minimally Invasive Surgery Training: A scoping review". Surgeries (Basel); 2025; 6(1):7.

[2] Pérez-Salazar MJ, Caballero D, *et al.*, "Correlation study and predictive modelling of ergonomic parameters in robotic-assisted laparoscopic surgery" Sensors (Basel); 2024; 24:e7721.

[3] Hermens HJ, Freriks B, *et al.*, "Development of recommendations for SEMG sensor and sensor placement procedures" J Electro Kines; 2000; 10:361-74.

[4] Dufaug A, Barthod C, *et al.*, "New joint analysis of electromyography spectrum and amplitude-based methods towards real-time muscular fatigue evaluation during a simulated surgical procedure: A pilot analysis on the statistical significance" Med Eng & Phys; 2020; 79:1-9.

A Transparent Simulation-Centered Co-Robotic Platform for Human-Supervised Spinal Surgery

Daniel Shabanianalavi[1], Pedro Machado[2], Kartikeya Walia[1] and Philip Breedon[1]

[1] *Medical Engineering Design Research Group, Nottingham Trent University, UK*

[2] *Computer Science Department, Nottingham Trent University, UK*

Corresponding author: philip.breedon@ntu.ac.uk

INTRODUCTION

Spinal procedures require precise planning, trajectory control, and continuous surgeon awareness, particularly in tasks such as drilling and pedicle screw placement, where small deviations may lead to neurological or vascular complications [1]. Robotic and AI-assisted systems are increasingly used to improve accuracy and reduce radiation exposure [2]. However, progress toward higher autonomy in surgery remains limited by a lack of transparency in robot behaviour and by the difficulty surgeons face in understanding what the robot is planning to do before and during a procedure [3–6]. This work addresses that gap through a co-robotic platform designed to make robot planning and workflow progression visible and interpretable to the surgeon. Importantly, the present study reports the current technical foundation created to support future trust-centered autonomy research. The main contribution is an integrated pipeline in ROS 2 Jazzy that combines MoveIt 2 motion planning, Gazebo-based surgical simulation, py_trees workflow visualisation, and a large-language-model (LLM) interface that translates robot state and surgery data into surgeon-friendly summaries. The system has the potential to enhance surgical perception by improving transparency before and during surgery.

MATERIALS AND METHODS

The implemented system is based on a ROS 2 Jazzy package connected to a Gazebo simulation environment and deployed on a dual-arm co-robotic spinal platform (Fig. 1). Within this environment, robot planning is executed using MoveIt 2, allowing the surgeon to preview planned robot motions on the simulated scene before execution on the real robot (Fig. 2) and to request a visualisation of these actions during surgery when needed. This creates a continuous visual representation of intended behaviour rather than presenting robot actions as opaque or unexpected events. To improve workflow transparency, the surgical process is represented using behaviour trees implemented in py_trees. The behaviour tree structure shows the current task node being executed, the relationship between sequential actions, and the next actions expected after task completion. An explanatory interface was also developed to connect multiple information sources, including robot hardware variables, arm positions, drill speed and torque, environment state, and behaviour-tree execution status, into a surgeon-facing display (Fig. 3). This data is passed to an LLM-based summarisation layer that produces surgeon-friendly explanations of the current robot and procedure state. The broader trust-centered information flow linking surgeon interaction, simulation, trust modelling, and robot action is illustrated conceptually in Fig. 4. At the current stage, this work constitutes an enabling architecture for future trust-model development.



Figure 1: Dual-arm co-robotic spinal surgery platform with two robotic manipulators positioned around a supine spinal phantom on the operating table.

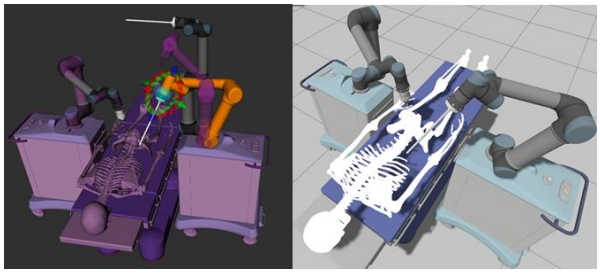


Figure 2: Gazebo and MoveIt 2 simulation scene showing the patient model, dual-arm robot configuration, and pre-execution motion-planning environment.

RESULTS

The current system provides a functioning prototype pipeline that integrates the physical robotic testbed, simulation-based motion planning, workflow representation, and explanatory feedback within one human-supervised robotic environment. The dual-arm co-robotic setup and spinal phantom arrangement are shown in Fig. 1. A ROS 2 Jazzy package was created and connected to Gazebo, enabling robot planning to be executed in a simulated surgical environment before execution on the real robot, as illustrated in Fig. 2. MoveIt 2 was successfully used to generate and visualise planned robot motions, while a py trees-based workflow layer exposed current actions, completed actions, and subsequent planned actions within the procedure. The developed interface aggregates robot and simulation variables and uses an LLM to produce readable summaries of the current system state for the surgeon (Fig. 3). Together with the trust-centered architectural concept shown in Fig. 4, these results demonstrate that the platform can already provide pre-operative and intra-operative visibility of robot planning and task progression, thereby increasing transparency of intended robot behaviour. At the same time, the study remains at a technical integration stage. The work provides the practical infrastructure required for future evaluation of trust, usability, workload, and human oversight in spinal robotic surgery.

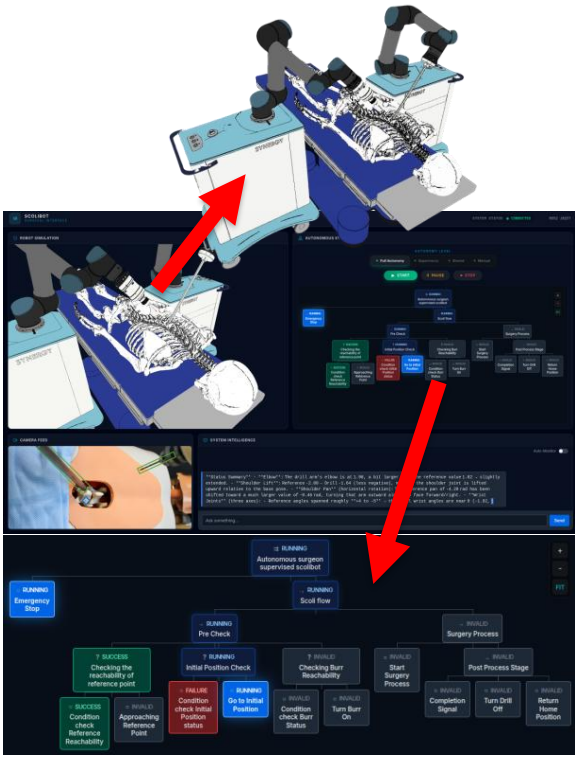


Figure 3: Integrated surgeon-facing interface showing robot simulation, autonomy state, live camera feed, and LLM-based system summary.

CONCLUSION AND DISCUSSION

This research represents the current technical stage of a wider PhD project on trust-centered co-robotic spinal surgery. The contribution is a transparent simulation and explanation pipeline, not yet a completed trust model. By combining Gazebo simulation, MoveIt 2 planning, behaviour-tree workflow exposure, and LLM-based state summarisation, the platform helps surgeons to see what the robot is doing, what it will do next, and how procedural context can be communicated in a more interpretable way. The main limitation is that the work is presently focused on system integration and transparency support rather than formal validation of trust or autonomy adaptation. Future work will build on this foundation to develop and evaluate a trust model, conduct surgeon-facing studies, quantify usability and workload, and investigate how explanation quality and workflow visibility influence confidence in robot-assisted spinal procedures.

REFERENCES

[1] Li X, Chen J, Wang B, Liu X, Jiang S, Li Z, Li W, Li Z, Wei F. Evaluating the Status and Promising Potential of Robotic Spinal Surgery Systems. Orthopaedic Surgery. 2024;16(11):2620–2632.

[2] Davidar AD, Jiang K, Weber-Levine C, Bhimreddy M, Theodore N. Advancements in robotic-assisted spine surgery. Neurosurgery Clinics. 2024;35(2):263–272.

[3] Sullins JP. Ethical trust in the context of robot assisted surgery. APA Newsletter on Philosophy and Computers. 2014;14(1):4–13.

[4] Haidegger T. Autonomy for surgical robots: Concepts and paradigms. IEEE Transactions on Medical Robotics and Bionics. 2019;1(2):65–76.

[5] Ficuciello F, Hamedani MH, Tamburrini G. Meaningful human control of increasingly autonomous robots for surgery. In: Research Handbook on Meaningful Human Control of Artificial Intelligence Systems. Edward Elgar Publishing; 2024:205–231.

[6] Lee A, Baker TS, Bederson JB, Rapoport BI. Levels of autonomy in FDA-cleared surgical robots: a systematic review. NPJ Digital Medicine. 2024;7(1):103.

[7] Nguyen CC, Wong S, Thai MT, Hoang TT, Phan PT, Davies J, Wu L, Tsai D, Phan HP, Lovell NH, Do TN. Advanced user interfaces for teleoperated surgical robotic systems. Advanced Sensor Research. 2023;2(4):2200036.

Development of a Pancreatic Phantom for Surgical Training

A. Callera[1,2], G. De Cristofaro[1,2], F. Potere[1], F. Cellesi[1], M. Pagnanelli[2,3], V. Santagati[2,3], G. Nappo[2,3], G. Cappretti[2,3], S. Carrara[4], A. Zerbi[2,3], M.L Costantino[1], and F. De Gaetano[1,2]

[1]*Department of Chemistry, Materials and Chemical Engineering "Giulio Natta", Politecnico di Milano, Italy*

[2]*IRCCS Humanitas Research Hospital, Italy*

[3]*Department of Biomedical Sciences, Humanitas University, Italy*

[4]*Department of Gastroenterology, Endoscopy Unit, IRCCS, Humanitas Research Hospital, Italy*

INTRODUCTION

Pancreaticoduodenectomy (PD) is one of the most complex and high-risk operations in abdominal surgeries. Postoperative pancreatic fistula (POPF), caused by leakage at the pancreatoenteric anastomosis, is a major source of postoperative complications. Although surgical techniques and perioperative care have evolved, outcomes after PD continue to depend heavily on the chosen surgical approach and the surgeon's experience. The learning curve for PD generally progresses through three phases: acquisition of competence, proficiency, and mastery. The number of procedures required to achieve proficiency for open PD is approximately 30 (20-50) procedures. Improvements in intraoperative parameters (operative time, blood loss) occur at the beginning of the second phase, while reductions in postoperative complications, such as POPF, emerge later with the reaching of surgical mastery. [1] To shorten this learning curve, developing an accurate training device, a surgical pancreatic phantom, is essential. Such a device would moreover allow for testing new techniques. In order to be effective it must replicate the mechanical properties of the natural organ. To this end, silicones and hydrogels are investigated here, as both have been reported for phantom production [2, 3]. In order to correctly replicate the surgery, the biological tissues involved during the procedure, namely pancreatic parenchyma, main pancreatic duct (MPD) and jejunum, are characterized.

MATERIALS AND METHODS

The mechanical properties of both biological and synthetic materials are analysed to develop a realistic surgical training phantom. Synthetic samples are tested following the same protocols as biological tissues to ensure a direct comparison.

In this work, mechanical tests are performed using instruments from Anton Paar GmbH (Graz, Austria). Indentation tests are conducted with the Bioindenter UNHT3, while compression, suture retention, and ring tests are carried out using the MCR 702e MultiDrive rheometer. All testing protocols are custom-made. Specifically, indentation, compression, and suture retention tests are conducted on the pancreas, tensile ring tests on the duct, and compression and suture retention tests on the jejunum.

To reproduce the pancreatic parenchyma, gellan gum is selected for its tuneable mechanical properties, biocompatibility, and realistic tactile feedback. Additionally, its hydrated appearance and viscoelastic properties make it suitable for mimicking soft biological tissues. To replicate the mechanical behaviour of the pancreatic duct, different materials and fabrication techniques are tested including silicones (DragonSkin, Smooth On), elastic 3D printing materials (Silicone 40A, Formlabs) and bioprinted hydrogels (gelatine-alginate). For the jejunum different silicones formulations (EcoFlex and DragonSkin, Smooth On) are considered.

The pancreatic phantoms developed in this study are evaluated by experienced surgeons from different centres across Italy to assess their validity in replicating the pancreatojejunostomy procedure. Each surgeon performed the anastomosis on two phantoms with different consistencies, alternating the order of testing. Afterward, they completed a questionnaire evaluating various aspects of the models, including suturability, tactile feedback, and suture-holding strength of both the parenchyma and the duct. The collected responses are analysed using the System Usability Scale (SUS) to quantify the phantom's usability. Additionally, information on the perceived usefulness of the device were gathered.

RESULTS

For the selection of phantom material for each tissue the most relevant property was identified. The results of gellan gum are shown in Figure 1.

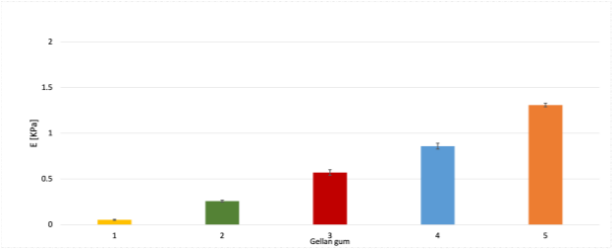


Figure. 1. Suture retention strength of tested materials for the pancreatic parenchyma.

The suture retention strength of pancreatic parenchyma is 2.96 ± 1.71N. The comparison between phantom

materials and biological duct and jejunum is shown in Figures 2-3.

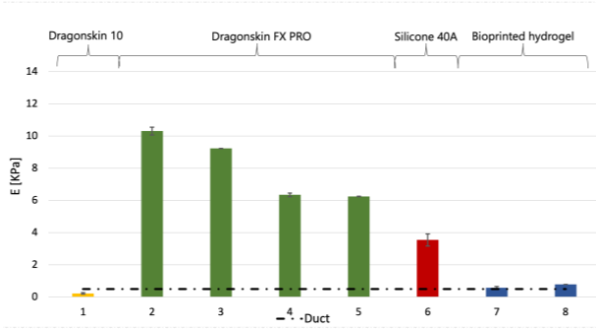


Figure. 2. Strain at failure of tested materials in comparison with the main pancreatic duct under tensile testing.

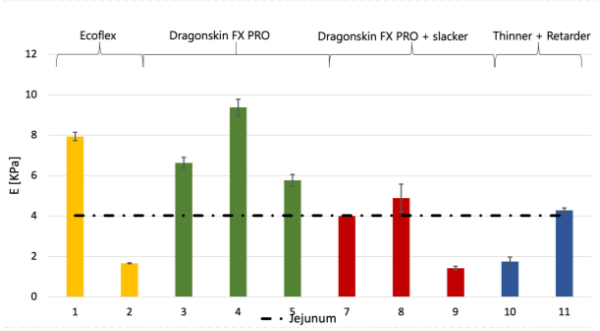


Figure. 3. Suture retention strength of tested materials in comparison with the jejunum.

The final assembled phantom is shown in Figure 4. The pancreatic phantom is composed of a dyed gellan gum, in compositions 4 and 5, 15 x 40 x 35 mm sample with a bioprinted hydrogel conduit, in composition 8, having a 5 mm diameter and 0.5 mm thickness. The jejunum is represented by a silicone tube, in composition 10, with 25 mm diameter and a 2 mm thickness.

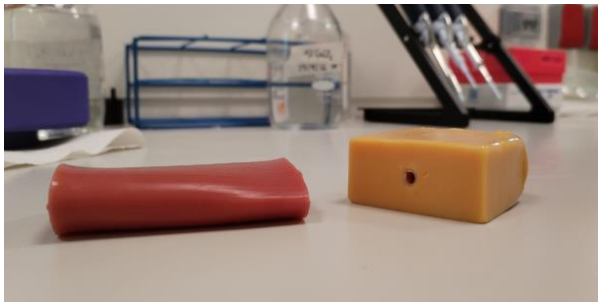


Figure. 4. Phantom for DCP including pancreatic parenchyma, Wirsung duct and jejunum

Values of SUS score varied between the two devices with phantom 2 tested second reaching 52%. The lowest scoring categories were the suture passage and holding strength of the duct and the interaction between duct and parenchyma. The highest ones were instead related to the consistency and suture holding strength of the parenchyma. The overall realism also consistently scored above 3 out of 5.

The majority of the interviewed surgeons found the device useful for training purposes as shown in Figure 5.

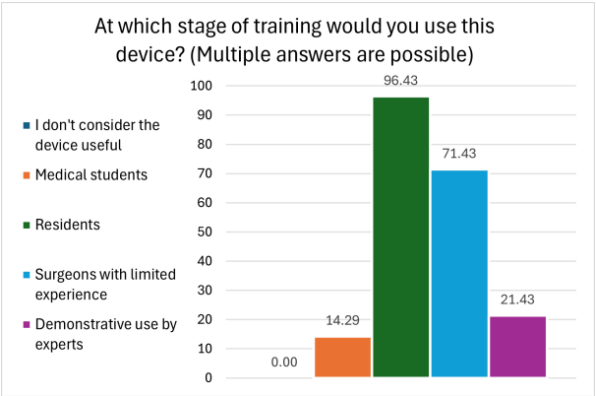


Figure. 5. Results of the question based on usefulness of the device in the final questionnaire.

CONCLUSION AND DISCUSSION

The mechanical characterization of biological tissues allowed the identification of key parameters describing pancreatic consistency.
In this study, a first version of the pancreatic phantom is developed using gellan gum for the parenchyma, hydrogel for the MPD, and silicone for the jejunum. Surgeons evaluated the prototypes and identified phantom 2 as the most promising for reproducing pancreatic consistency, finding it useful for the residents' training. However, improvements are needed, particularly regarding the duct's mechanical response during suturing and the limited interaction between the phantom components.
After the improvement of the device the next step will be to test it with residents multiple times in order to observe whether or not it is able to improve technique and show a learning curve trend.

REFERENCES

[1] Joshi K, Espino DM, Shepherd DE, Mahmoodi N, Roberts KJ, Chatzizacharias N, Marudanayagam R, Sutcliffe RP. "Pancreatic anastomosis training models: Current status and future directions." Pancreatology. 2024 Jun;24(4):624-629.

[2] Wang Y, Tai B L, Yu H, Shih A J. "Silicone- Based Tissue-Mimicking Phantom for Needle Insertion Simulation." ASME. J. Med. Devices. June 2014; 8(2): 021001.

[3] Chen RK, Shih AJ. "Multi-modality gellan gum-based tissue-mimicking phantom with targeted mechanical, electrical, and thermal properties. " Phys Med Biol. 2013 Aug 21;58(16):5511-

Intuitive Robot-Assisted Drilling with Haptic Feedback and Multi-Layer Safety Constraints

Roman Leonov[1], Johannes Hölker[1], Do Yeon Kim[1], Dennis Kundrat[1,2], and Jan Erik Hermann Wolff[1,3]

[1] *Fraunhofer Research Institution for Individualized Medical Technology and Engineering, Lübeck, Germany*
[2] *Lübeck University of Applied Sciences, Lübeck, Germany*
[3] *University Hospital Schleswig-Holstein, Department of Oral and Maxillofacial Surgery, Lübeck, Germany*

INTRODUCTION

Orthopedic and Craniomaxillofacial (CMF) surgery requires precise placement of osteosynthesis plates, screws and patient-specific implants, for which bone drilling is the first and decisive step. Because major vessels and peripheral nerves lie close to the operative targets, even minor positional errors can cause irreversible nerve or vascular injury, and excessive drilling heat can lead to bone necrosis. Robotic assistance has the potential to improve both accuracy and safety: platforms have improved arthroplasty (MAKO, Stryker; ROSA Knee, Zimmer Biomet; NAVIO, Smith & Nephew) [1], [2], pedicle screw placement (Mazor X, Medtronic; ExcelsiusGPS, Globus Medical) [3] and stereotactic neurosurgery [4]. All are designed for specific anatomical regions and high-load conditions, and none offers the dexterity and compact form factor required in the confined operative fields of CMF surgery. This work presents a teleoperated drilling system, its multi-layer safety framework, and first phantom measurements of drilling accuracy.

MATERIALS AND METHODS

A. Robotic Setup

An industrial drilling spindle (Mechatron HFSAC-6508-24-ER11; Mechatron GmbH, Germany) is attached as the end-effector to a Kuka LBR Med 14 (KUKA AG, Augsburg, Germany). The spindle provides a rated speed of 24,000 rpm, within the range reported for high-speed surgical drill systems in orthopaedic procedures [5]. The control system is built in a Robot Operating System 2 (ROS2) environment using the LBR-stack by [6]. The surgeon manually controls the movement using a Sigma 7 haptic device (Force Dimension, Nyon, Switzerland). The proposed setup is demonstrated in Figure 1.

We decouple the spatial control of the end-effector into linear and angular movements. The surgeon can therefore navigate to a position with a fixed spindle axis and set the spindle orientation with a fixed drill-bit tip position in separate steps, simplifying the drilling of contiguous parallel holes, a common requirement in plate-fixation procedures.

B. Safety Feature Implementation

The primary kinematic layer restricts motion during drilling to the drill-bit axis, with the feed rate capped at 1.0 mm/s [7] and a user-defined depth limit. Two mechanisms address tissue-specific hazards. First, cumulative drill time is monitored against a threshold derived from the 47 °C / 60 s necrosis limit [8]; on reaching it the manipulator retracts with reduced, reversed rotation. This sequence cannot be overridden and can also be triggered by a brief pullback. Cumulative time is a conservative but indirect proxy, as the temperature reached also depends on feed rate, drill geometry, irrigation and bone density; validation is planned. Second, breakthrough detection integrates depth from the commanded axial velocity and suppresses forward feed when advance increases abruptly; detection latency and overshoot remain to be quantified.

Since these constraints reduce the operator's degrees of freedom and thus perceived control [9], two cues are rendered on separate channels to avoid perceptual interference: a spring-like restoring force proportional to the deviation of the hand from the permitted motion subspace (within the device limits of 20 N and 0.4 Nm), and a 220 Hz vibration superimposed once spindle current exceeds the no-load baseline, matching the peak sensitivity of the Pacinian corpuscles [10].

An eye-tracking system (Tobii Pro Glasses 3; Tobii AB, Stockholm, Sweden; gaze accuracy 0.6°, 100 Hz) provides an additional supervisory safety layer preventing motion when the surgeon's attention leaves the operative field. An



Fig. 1: Test setup for orthopedic drilling with a spindle attached to a Kuka LBR Med14 and a Force Dimension Sigma 7 motion tracking device

ArUco marker [11] on the end-effector (Fig. 1b) defines a circular area of interest (AOI) in the scene camera image of an eye tracker (Tobii Pro Glasses 3; 0.6°, 100 Hz), and commands are transmitted only while the gaze point lies inside it. On gaze departure, marker loss or occlusion the system fails safe as above: the spindle keeps rotating and the manipulator withdraws to a defined clearance, so the bit is never left embedded in bone. An ongoing retraction is exempt from gating and cannot be overridden.

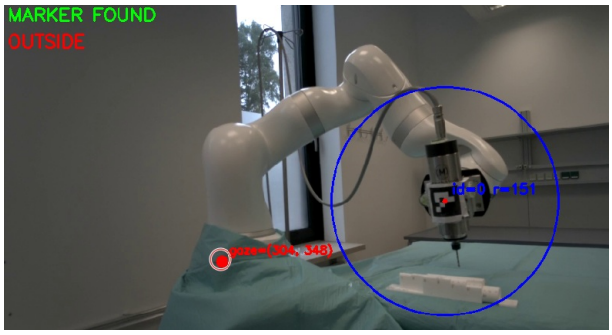


Fig. 2: Manipulator with spindle, drill bit and marker during the drilling process of a test object

C. Phantom Accuracy Evaluation

Drilling accuracy was assessed on polyamide (PA 12) specimens comprising seven regions with three holes each (21 holes per condition), covering perpendicular and 45° drilling on flat, convex and inclined surfaces. Specimens were scanned by micro-CT (Yxlon FF35CT) and evaluated by least-squares cylinder fitting. Robot-assisted drilling with an operator-defined drill axis was compared against freehand drilling performed by the same experienced surgeon.

RESULTS

Inter-hole angular offset (parallelism) and cylindricity were better in the robot-assisted case (Table I). The weakest robotic result occurred on a convex surface (10 mm radius, $11.02 \pm 6.23°$), where the constant feed rate advanced the bit during lateral slippage. These data are descriptive: a single operator performed both the freehand and the robot-assisted condition in one session, no statistical testing was applied, and PA 12 does not reproduce the properties of a bone.

TABLE I: Drilling accuracy on PA 12 phantoms (mean ± SD, $n = 21$ holes per condition).

Metric	Freehand	Robotic
Relative axis dev. (°)	4.10 ± 2.92	$\mathbf{3.85 \pm 3.81}$
Absolute axis dev. (°)	$\mathbf{9.44 \pm 4.62}$	9.84 ± 11.01
Cylindricity (μm)	185 ± 26	$\mathbf{178 \pm 22}$

CONCLUSIONS AND DISCUSSION

Each safety layer of the presented system targets a distinct failure mode: thermal injury, cortical breakthrough, inadvertent motion and loss of surgeon attention, while the haptic interface preserves the intuitive feel of manual drilling. A first phantom study indicates accuracy comparable to freehand drilling by an experienced surgeon. Validation of the safety layers themselves, breakthrough latency and overshoot, thermal profiles, and gaze gating under occlusion, is ongoing on bone-equivalent surrogates and cadaveric specimens.

REFERENCES

[1] E. L. Hampp, M. Chughtai, L. Y. Scholl, N. Sodhi, M. Bhowmik-Stoker, D. J. Jacofsky, and M. A. Mont, "Robotic-arm assisted total knee arthroplasty demonstrated greater accuracy and precision to plan compared with manual techniques," *Journal of Knee Surgery*, vol. 32, no. 3, pp. 239–250, 2019.

[2] A. Seidenstein, M. Birmingham, A. Plotkowski, and G. J. Golladay, "Better outcomes with robotic compared to manual unicompartmental knee arthroplasty: a systematic review," *Knee Surgery, Sports Traumatology, Arthroscopy*, vol. 29, no. 3, pp. 859–866, 2021.

[3] S. R. Kantelhardt, R. Martinez, S. Baerwinkel, R. Burger, A. Giese, and V. Rohde, "Perioperative course and accuracy of screw positioning in conventional, open robotic-guided and percutaneous robotic-guided pedicle screw placement," *European Spine Journal*, vol. 20, no. 6, pp. 860–868, 2011.

[4] M. Lefranc, C. Capel, A.-S. Pruvot-Occean, A. Fichten, C. Desenclos, P. Toussaint, D. Le Gars, and J. Peltier, "Frameless robotic stereotactic biopsies: a consecutive series of 100 cases," *Journal of Neurosurgery*, vol. 122, no. 2, pp. 342–352, 2015.

[5] G. Augustin, T. Zigman, S. Davila, T. Udilljak, T. Staroveski, D. Brezak, and S. Babic, "Cortical bone drilling and thermal osteonecrosis," *Clinical Biomechanics*, vol. 27, no. 4, pp. 313–325, 2012.

[6] M. Huber, C. E. Mower, S. Ourselin, T. Vercauteren, and C. Bergeles, "LBR-Stack: ROS 2 and Python Integration of KUKA FRI forMed and IIWA Robots," *Journal of Open Source Software*, vol. 9, no. 103, p. 6138, Nov. 2024.

[7] B. Zhang, G. Chen, Y. Lv, H. Li, and H. Wei, "Methods of bone drilling in surgery: current situation and prospects," *The International Journal of Advanced Manufacturing Technology*, vol. 137, no. 1-2, pp. 23–45, Mar. 2025.

[8] R. A. Eriksson and T. Albrektsson, "The effect of heat on bone regeneration: An experimental study in the rabbit using the bone growth chamber," *Journal of Oral and Maxillofacial Surgery*, vol. 42, no. 11, pp. 705–711, 1984.

[9] P. Schleer, S. Drobinsky, M. De La Fuente, and K. Radermacher, "Toward versatile cooperative surgical robotics: a review and future challenges," *International Journal of Computer Assisted Radiology and Surgery*, vol. 14, no. 10, pp. 1673–1686, Oct. 2019. [Online]. Available: http://link.springer.com/10.1007/s11548-019-01927-z

[10] S. J. Bolanowski, G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four channels mediate the mechanical aspects of touch," *The Journal of the Acoustical Society of America*, vol. 84, no. 5, pp. 1680–1694, 1988.

[11] F. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, 06 2018.

Kinematic considerations in laparoscopic surgical robots. Impact on usability

Josep Amat[1,2], Alícia Casals[1,2], Narcís Sayols[1]

RobSurgical Systems [1]

Universitat Politècnica de Catalunya (Emeritus) [2]

alicia.casals@upc.edu

INTRODUCTION

Robotic laparoscopic surgery was introduced in operating rooms in the early 2000s by Intuitive Surgical with the DaVinci robot, which over about 20 years has been the only robot on the market and is currently the leader in this surgery.

The challenge of robotic laparoscopic surgery is to be able to reproduce with the instruments the movements that a surgeon performs from the teleoperation console, since the operation is still guided by the surgeon himself. Autonomous surgery, with AI, is still a promise, since it only achieves some incipient positive results in research. In both cases, teleoperated or when AI makes autonomy possible, the robot must be able to move 4 arms without collisions or blockages. Redundancy in this robotic surgery is necessary to ensure operability with multiple arms without collisions and to be able to place the arms in different configurations operating in the same surgical space.

In recent years, other robotic systems have appeared, taking advantage of the expiration of many of the multiple patents of this pioneering robot. These robots have followed different variants, either multi-arm robotic systems, or single arm multi-robot systems. The robots in each of these two groups differ by the high or low degree of redundancy that the system offers to avoid collisions between arms.

DaVinci has marked a first family of robotic systems, that of a multi-arm robot with high redundancy. This same structure has been followed by other companies, especially Chinese [1, 2]. Less redundancy is offered by Bitrack, our system, also a multi-arm system [3]. This system, the result of a spin-off from the Technical University of Catalonia, started in 2016, has recently been validated for urology with the CE mark, and is currently in the phase of commercial introduction and extension to general surgery. A second family is made up of multi-robot systems, with arms also with different degrees of redundancy. Those with high-redundancy robot arms follow an architecture equivalent to that of the DaVinci arms.

In both, multi-robot and multi-arm robot systems, there is a wide range of redundancy levels, from systems with up to 5 redundant DoFs to systems with only 1.

MATERIALS AND METHODS

In robotic surgery, each arm that supports an instrument must have the necessary degrees of freedom to position the instrument (3DoF) and to orient it (2DoF, pitch and yaw). Having 5DoF on each arm will therefore allow to operate, but it is necessary to avoid interference between arms when working so close to each other in the surgical field. For this reason, it is necessary to introduce redundant degrees of freedom to allow positioning the arms in different ways during the operation. Figure 1 shows the different alternatives in number of redundant DoFs introduced.

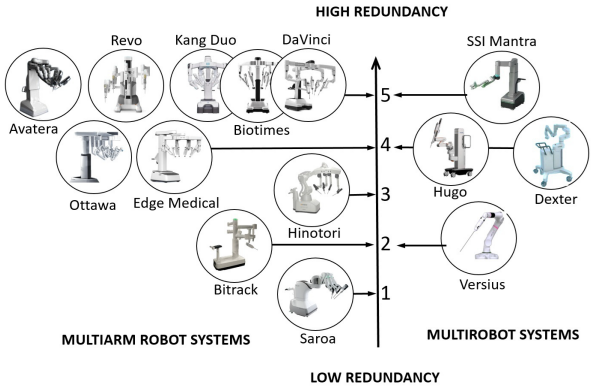


Figure. 1 Alternatives of laparoscopic robotics systems

To ensure that collision situations are solved, DaVinci introduced from the beginning 5 additional joints that can provide 5 redundant DoF, without considering the P and Q joints of the entire head containing the 4 arms, fig.2.

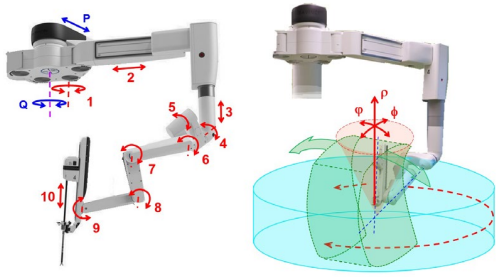


Figure. 2 Different levels of control in a highly redundant polyarticulated structure

In this system, the architecture is designed to operate at two levels. In a first level of manual control, a structure of 3 proximal joints, W1, W2 and W3 of Scara architecture (figure 2), together with the common P and Q joints, allows the positioning of a second structure within a three-dimensional space of cylindrical coordinates, in blue in figure 2. This second structure of 7 joints allows to approximate and fix its distal point on the trocar with 3 DoF within the volume (in green) by

means of the joints W4-W9. During the operation, the instrument must pivot on the RCM (Remote Center of Motion), the position of the trocar, in a space of polar coordinates of 3 DoF (in red) by means of the joints W6-W9, which share the function of positioning on the trocar. Finally, joint W10 controls the penetration within this polar space. Therefore, this arm has 1 additional effective redundant DoF for pose. The advantage of having such a polyarticulated structure is that the movements of the robot during the operation are carried out by the distal joints, thus facilitating the precise control of the robot. This high redundancy also entails, in addition to greater complexity and cost, a greater physical occupation of space above the patient, which makes it difficult for medical personnel to access the patient when necessary.

Other more recent systems have followed this same pattern, high redundancy, either with multi-robot or multi-arm systems, forming the group of high redundancy, as shown in Figure 1.

High redundancy in multi-arm systems entails wide possibilities of initial configuration of each arm. In multirobot systems, there is also the possibility of placing each robot in the most convenient place, which represents an additional redundancy.

This high redundancy availability, by itself, does not resolve all possible conflict situations, and implies the need for some experience in defining the initial configuration of the four arms to reduce the long set-up time invested in verifying operation without collisions before starting an operation.

At the other extreme are systems in which the degree of redundancy has been reduced to two, or even a single DoF, Fig.1. In this case, low redundancy can mean that the optimal pose can be determined by the robot control unit itself, so that this redundancy is transparent and therefore unnoticed by the user, which greatly simplifies the set-up. These more simplified structures usually also mean a smaller workspace occupancy.

Figure 3 shows the redundancy of two low redundancy systems. In the first case, Bitrack, in which the three-dimensional positioning of the tool is achieved with the joints W1, W2 and W3, in an angular coordinate space. The movements of the instrument that must pivot on the position of the trocar, in a polar coordinate space of 3 DoF (in red) are obtained combining the movement of the joints W1-W6. In this case the two joints W5 and W6 of the two lower arms (arms 1 and 2), having 1 redundant joint, are passive (in purple) to intrinsically ensure that the instrument pivots on the trocar without generating any physical effort on the patient. In the non-redundant upper arms (3 and 4) the passive joints are W4 and W5. In this system, an arm at risk of collision can move away from the other by acting on the redundant W4 joint of the lower arms, which allows them to move away with a single DoF within the vertical plane indicated in blue in the figure. Another redundant DoF, W0, is shared by all arms to adapt to the patient's position.

In the second case, the Versius robot [4] is a multirobot system in which each arm has a classic structure of angular coordinates, but with two additional joints, W3 and W5, which give the assembly a redundant DoF that allows the elbow of the W4 joint to move laterally, within the plane marked in blue in figure 3, maintaining the position of the instrument that must pivot on the position of the trocar in the polar workspace, marked in red.

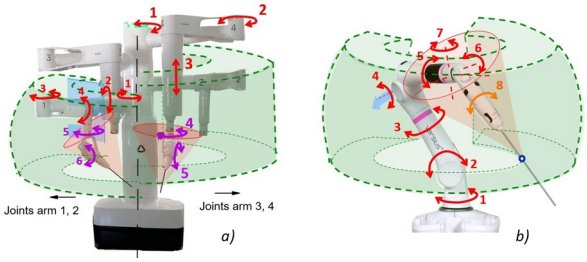


Figure. 3. Kinematic structure of (a) the Bitrack and (b) the Versius systems

RESULTS

The set-up times have been compared between a high redundancy robot, the DaVinci, and a low redundancy robot, the Bitrack. In the first case, according to previous published studies of multiple interventions, resulted in a range from 9 to 50 min and the average values are around 20 min. Progress with DaVinci Xi have now reached a mean time of 8 min [5]. In the second case, the times measured in the 20 interventions carried out in the clinical validation process for the CE marking approval of the Bitrack system, the results range from 1.8 to 8.5 minutes, with an average time around 3,5 minutes. The medical results, both in terms of the duration of the interventions and the postoperative complications, have not been found to differ between the different systems.

CONCLUSION AND DISCUSSION

We can conclude that the different systems ranging from high to low redundancy robots, both in multirobot or multi-arm systems offer different performances. The main differences offered by the Bitrack system are manifested in the shorter set-up time and the fact that it presents less difficulty in operating in the four quadrants. These factors represent an improvement not in medical effectiveness, but in efficiency. Regardless of the different footprint occupied by one or other systems.

REFERENCES

[1] Editorial. “New surgical robotic platforms in China and their applications in urologic surgeries“, Asian Journal of Urology, Science Direct, 2025.

[2] S. Niu et al. “Suitability of the MP1000 Platform for Robot-assisted Prostatectomy: A Prospective Randomised Controlled Trial“, European Urology Open Science, 2024

[3] J. Amat, A. Casals and M. Frigola. “Bitrack: a friendly four arms robot for laparoscopic surgery“, CRAS 2020

[4] C. Gaia et al. Versius surgical system: tips and tricks for OR setting and port placement for pelvic surgery, our experience in a multi-robotic referral centre. Langenbeck's Archives of Surgery, 2025

[5] Mohamed H, Khanna S, Mallika MG, Karuthedath N. A comparative study of draping time and docking time of da Vinci Xi and multiple independent arm robotic platforms. Int J Adv Robot Innov Surg 2024;2:82-6

Physics-Guided Residual Modeling for Transient Prediction in Pneumatic Soft Continuum Robots

Q. Wei[1], J. Liu[2], M. Chauhan[1,*]

[1]*School of Physics, Engineering and Technology, University of York, York, United Kingdom*
[2]*Chengdu University of Information Technology, Chengdu, China*
wsp525@york.ac.uk; geekliuowo@outlook.com; *manish.chauhan@york.ac.uk

INTRODUCTION

Pneumatic soft continuum robots show promise for medical manipulation and navigation through confined spaces [1], [2]. Under rapid actuation they overshoot and ring, so a steady-state pressure-angle map is not enough. Finite-element models are accurate but too slow for real-time use, and pure data-driven surrogates throw away physical interpretability. Others have combined them: Gao *et al.* trained residual physics to close a sim-to-real gap [3], Johnson *et al.* paired first principles with deep learning for model-based control [4], and Bruder *et al.* used a Koopman-based residual on a soft arm [5].

We propose a physics-guided residual model with three pieces. First, a differentiable physics block with seven parameters covering chamber asymmetry, quadratic force-pressure scaling, and a cubic large-bending correction. Second, a residual MLP initialized to zero. Third, trajectory matching for training, which keeps us off the noisy ground that derivative-target residual learning sits on. The combination beats a same-capacity pure-MLP baseline in simulation while keeping every parameter interpretable.

MATERIALS AND METHODS

Fig. 1 illustrates the proposed framework. We define the state as $\mathbf{x}_t = [\theta_t, \omega_t, p_{1,t}, p_{2,t}]^\top$ and the input as $\mathbf{u}_t = [p_{1,\mathrm{cmd},t}, p_{2,\mathrm{cmd},t}]^\top$, where $p_{i,t}$ and $p_{i,\mathrm{cmd},t}$ are the measured and commanded pressures of chamber i. Each chamber pressure follows a first-order response: $\dot{p}_{i,t} = \tau_i^{-1}(p_{i,\mathrm{cmd},t} - p_{i,t})$.

Differentiable physics: Rather than the usual symmetric textbook form, we use a richer seven-parameter structure motivated by soft-pneumatic considerations,

$$\begin{aligned}\dot{\omega}_t = {} & b_1 p_{1,t} + \alpha_1 p_{1,t}^2 - b_2 p_{2,t} - \alpha_2 p_{2,t}^2 \\ & - \beta_c \omega_t - \beta_k \theta_t - \kappa_3 \theta_t^3,\end{aligned} \tag{1}$$

where $b_1 \neq b_2$ encodes chamber asymmetry, $\alpha_i p_{i,t}^2$ follows the well-known quadratic force-pressure relation of pneumatic actuators [2], and $\kappa_3 \theta_t^3$ lumps together the geometric effects of large bending: changing moment arm, chamber rearrangement, and constraint-layer pull-back. For typical silicone elastomers in normal bending ranges, intrinsic material stiffening contributes little, so we treat κ_3 as a geometric rather than constitutive parameter. The full physics $f_{\mathrm{phys}}(\mathbf{x}_t, \mathbf{u}_t; \phi)$ is differentiable in $\phi = \{b_i, \alpha_i, \beta_c, \beta_k, \kappa_3, \tau_i\}$; positivity-constrained entries are stored in log-space.

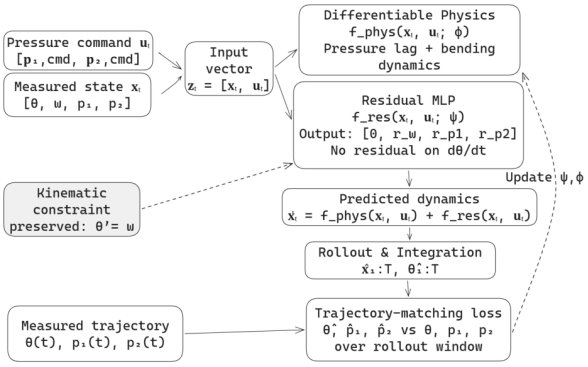


Fig. 1: Physics-guided residual framework.

Residual composition: Equation (1) misses several effects we know are present: inter-channel coupling, quadratic damping, valve hysteresis, and leakage. We let a residual network absorb them,

$$\dot{\mathbf{x}}_t = f_{\mathrm{phys}}(\mathbf{x}_t, \mathbf{u}_t; \phi) + f_{\mathrm{res}}(\mathbf{x}_t, \mathbf{u}_t; \psi), \tag{2}$$

where f_{res} is a three-layer tanh MLP with 64 hidden units, writing into $[\dot{\omega}, \dot{p}_1, \dot{p}_2]$ only; the kinematic identity $\dot{\theta}_t = \omega_t$ is hardcoded. The MLP's final layer is zero-initialized, so at epoch 0 the model is the warm-started physics block, and the residual stays small unless the data force it up.

Trajectory-matching training: Fitting a residual against a numerically-differentiated target $\dot{\mathbf{x}}^{\mathrm{est}} - f_{\mathrm{phys}}$ pushes second-order finite-difference noise into the learning signal. We avoid this with chunked trajectory matching: at each of W start points on the training segment, we roll (2) forward for T steps with explicit Euler and minimize

$$\mathcal{L}(\phi, \psi) = \frac{1}{WT} \sum_{w,j} \Big[(\theta_{w,j} - \tilde{\theta}_{w,j})^2 + \lambda_p \sum_i (p_{i,w,j} - \tilde{p}_{i,w,j})^2 \Big], \tag{3}$$

where $\tilde{\cdot}$ are measured signals and $\lambda_p = 0.3$ balances angle against pressure terms. We update ϕ and ψ jointly with Adam at learning rates 5×10^{-4} and 5×10^{-3} respectively, clipping gradients at norm 10.

RESULTS

Experimental Setup: The reference simulation deliberately steps outside (1): it adds inter-channel coupling, quadratic damping, and valve nonlinearities (saturation, hysteresis). Inputs are random piecewise-ramp pressure

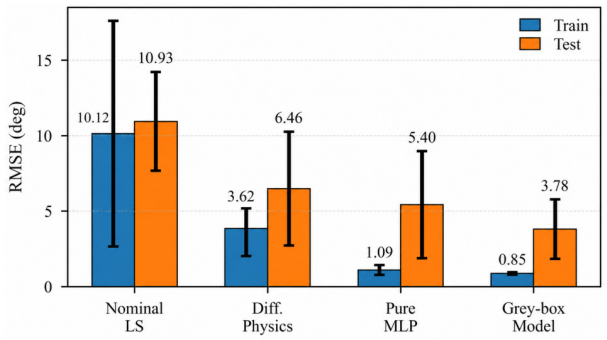


Fig. 2: Train and test RMSE over $N = 5$ trials (mean ± std). Our grey-box model has the lowest RMSE on both segments and a smaller train-to-test gap than the same-capacity pure MLP.

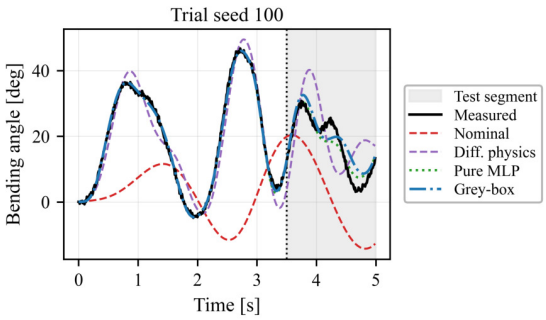


Fig. 3: Representative simulation trial (seed 100). Shaded: held-out test segment.

commands, and Gaussian measurement noise ($\sigma_\theta = 0.5°$, $\sigma_p = 0.02$ bar) is injected on every signal we feed downstream, so both parameter identification and neural training run on noisy data. Each trajectory is split 70/30 in time; we report mean and standard deviation over $N = 5$ independent trials, each with its own random seed for command and noise. Three baselines run alongside ours: a 3-parameter *nominal* model fit by least squares and frozen, a same-capacity *pure MLP* trained against the same loss, and a *differentiable physics only* ablation that trains (1) without any residual.

Performance: Table I reports per-segment RMSE and Fig. 2 shows the same numbers as paired bars. The 3-parameter nominal model is structurally inadequate (10.93° test RMSE). Letting the seven physics parameters train alone, with no neural component, already brings test RMSE down to 6.46°. A same-capacity pure MLP reaches 5.40°. Our grey-box model gets to 3.78°, a 1.62° improvement (30 % on average) that exceeds one standard deviation across trials. Fig. 3 shows a representative trial: on the held-out segment our grey-box prediction tracks the measured angle closely, while the pure MLP drifts and the nominal model is off in both phase and amplitude.

In practice, an RMSE of 3.78° means the predicted bending is off by about 4° on average. For a 100 mm soft actuator that translates to roughly 7 mm of tip-position error. It is enough to matter for path planning but not catastrophic. Reducing this error makes pressure planning more reliable and lowers the risk of under-bending, overshoot, or excessive pressurisation. The residual component therefore mainly captures unmodelled transient dynamics, such as inter-chamber coupling, valve and pneumatic-line nonlinearities, and velocity-dependent damping. We also note that train-to-test RMSE grows by 1.7× for the physics-only ablation and 4 to 5× once a neural component is present, the simulation-to-real risk that hardware validation will need to confront.

TABLE I: RMSE over $N = 5$ independent trials, each generated with a different random seed controlling both the pressure command sequence and the measurement-noise realization (mean ± std, degrees).

Model	$\text{RMSE}_{\text{train}}$	$\text{RMSE}_{\text{test}}$
Nominal (3-param LS, frozen)	10.12 ± 7.50	10.93 ± 3.30
Diff. Physics (7-param, no residual)	3.62 ± 1.80	6.46 ± 3.70
Pure MLP (no physics)	1.09 ± 0.40	5.40 ± 3.50
Physics-Guided Residual (ours)	**0.85 ± 0.05**	**3.78 ± 2.00**

CONCLUSION AND DISCUSSION

We developed a hybrid soft-robot trajectory predictor by integrating a seven-parameter differentiable physics block with a neural residual network. The physics block provides an interpretable structural, and the residual captures the dynamics outside the physics model.

On simulated data, the model cut test RMSE by 30 % against a same-capacity pure-MLP baseline, with every physical parameter inspectable. But train-to-test RMSE still grows by a factor of four to five once an MLP is in the loop, and only hardware testing will tell us how much of that survives in practice.

The model is end-to-end differentiable, so it provides analytic Jacobians that drop into a gradient-based model-predictive controller. That makes it a fit for safety-constrained applications such as minimally invasive navigation, where physical and motion bounds need to hold during planning. Future work will validate it on hardware and embed it in a predictive control loop.

REFERENCES

[1] E. Franco, T. Ayatullah, A. Sugiharto, A. Garriga-Casanovas, and V. Virdyawan, "Nonlinear energy-based control of soft continuum pneumatic manipulators," *Nonlinear Dyn.*, vol. 106, no. 1, pp. 229–253, 2021.

[2] M. S. Xavier, C. D. Tawk, A. Zolfagharian, J. Pinskier, D. Howard, T. Young, J. Lai, S. M. Harrison, Y. K. Yong, M. Bodaghi, and A. J. Fleming, "Soft pneumatic actuators: A review of design, fabrication, modeling, sensing, control and applications," *IEEE Access*, vol. 10, pp. 59442–59485, 2022.

[3] J. Gao, M. Y. Michelis, A. Spielberg, and R. K. Katzschmann, "Sim-to-real of soft robots with learned residual physics," *IEEE Robot. Autom. Lett.*, vol. 9, no. 10, pp. 8523–8530, 2024.

[4] C. C. Johnson, T. Quackenbush, T. Sorensen, D. Wingate, and M. D. Killpack, "Using first principles for deep learning and model-based control of soft robots," *Front. Robot. AI*, vol. 8, 654398, 2021.

[5] D. Bruder, D. Bombara, and R. J. Wood, "A Koopman-based residual modeling approach for the control of a soft robot arm," *Int. J. Robot. Res.*, vol. 44, no. 3, pp. 388–406, 2025.

Quantitative Magnetic Centralisation for Continuum Robot MRI Localisation

Yael L. May, Peter Lloyd, James H. Chandler, Pietro Valdastri and Jürgen E. Schneider

sc18ylm@leeds.ac.uk, STORM Lab, IRASS, School of Electronic and Electrical Engineering, University of Leeds

INTRODUCTION

Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality without ionising radiation used to aid medical diagnosis and therapeutic treatments. Furthermore, MRI can be utilised not only to localize medical soft robots but also to control them via magnetic fields inherently available on an MR system [1], [2], [3].

One of the largest limitations of this combination is that the required ferromagnetic components inside the sensitive imaging volume cause a signal void, that can be 50 to 100 times the size of the object [1]. This makes localisation and navigation particularly challenging. Ongoing research focuses on reducing signal void artefacts by employing bespoke hardware solutions or sequences and minimising ferromagnetic parts [2], [4], [5]. These are typically ferromagnetic beads [1], [3].

A common assumption in magnetic MRI-driven robotics research is that the ferromagnetic ball bearing is located at the geometric centre of the signal void. This work proposes a localisation framework that combines dipole computational modelling with a centre-localisation algorithm applied to experimental data acquired from a preclinical high-field MRI system. The results demonstrate accurate localisation of a 1 mm low carbon steel (LCS) ball bearing and ferromagnetically tipped continuum robot derived from MR images, the most common approach for actuation in MRI robotics, despite a substantial signal void.

MATERIALS AND METHODS

Imaging in all 3 dimensions (3D) contributes greater object position accuracy, this can then be broken down into a series of 2D and 1D representations [3]. For this work, data was acquired on a Bruker 7 T BioSpec 70/20 MR system using a 3D Gradient Echo sequence (matrix size: 128 x 128 x 128). In Fig. 1, the samples scanned used a 1 mm LCS ball bearing in a cavity at the centre of a 25 mm diameter 3D crosshair (Fig. 1c), inside an acrylic sphere of diameter 32 mm. The sphere was filled 2% Agarose in water with a $2\mu M$ Dotarem Gadoteric Acid MRI contrast.

The experimental data was processed using custom Python software that reconstructed the raw data into a three-dimensional cubic dataset. To generate the experimental signal shown in Fig. 2b, each row and column corresponding to signal traversing the ball bearing, as in a millirobot, was extracted and reconstructed as a 1D phase map. To gain the 3D centre point, the dataset was analysed via 3D slicing through the geometric centre of the domain. A binary mask was applied to each slice to identify the edge boundaries of

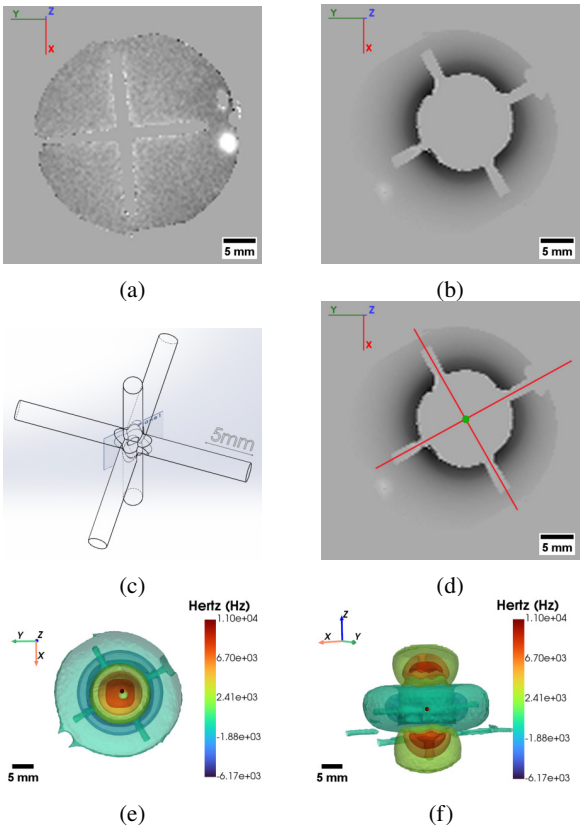
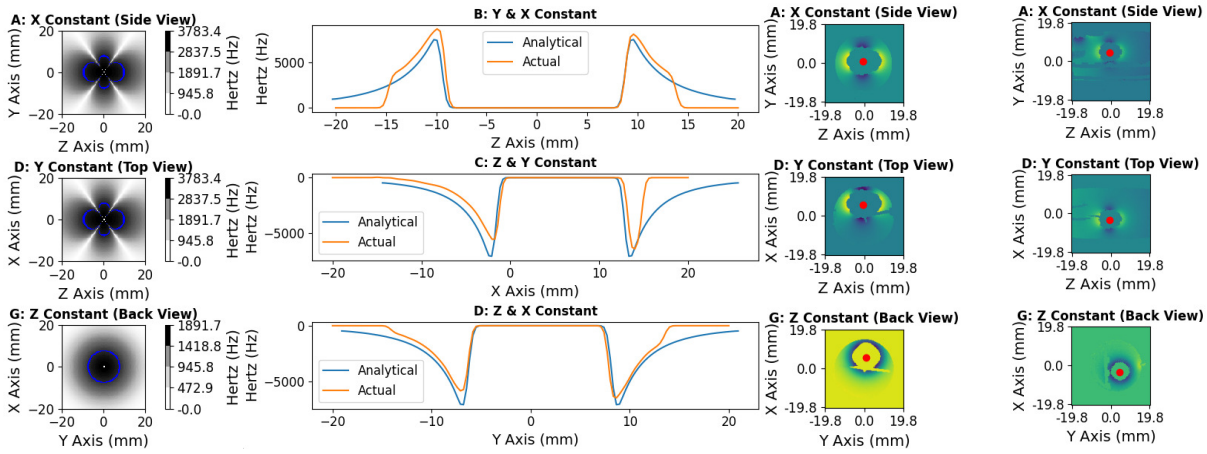


Fig. 1: (a) A 32 mm plastic sphere with a resin crosshair and 1 mm resin ball bearing in the centre and filled with an agarose-contrast solution. (b) An example of the sphere with a crosshair with a 1 mm LCS ball bearing. (c) SolidWorks design of the 25 mm diameter crosshair inside the MRI samples. (d) Image reconstruction showing the purpose of the crosshair with the LCS ball bearing in the centre (shown in green). (e) Mid slice out of a 3D stack (in Hertz) with a 1 mm LCS ball bearing (shown in dark red). (f) Full 3D reconstruction of the experimental signal void (in Hertz) due to field perturbations caused by the 1 mm LCS ball bearing (shown in dark red).

the signal void, defined as the region of zero signal intensity. When two distinct edges of the void were detected, the midpoint was found in the corresponding trough subset to estimate the centre of the signal void.

In the absence of a discernible trough, when the signal void was located far from the geometric centre, a connected neighbour average of more than 10 pixels was used instead. The estimated centre was iteratively refined through five recursive iterations to achieve convergence on a central value. A 1 mm red dot was plotted on each 2D reconstruction image to show where the calculated centre is, as in Fig. 2c.

To verify that the signal void produced by the ball bearing originated at the centre of the experimental data and that



(a) (b) (c) (d)

Fig. 2: (a) MRI Analytical Model with field maps. (b) Overlapping results between the analytical model and experimental data. (c) Example experimental data as phase maps and 2D images. (d) Centering algorithm applied to a continuum robot to get the centre [-3.125, 4.6875, -0.3125] in millimetres.

the identified centre was accurate, a 3D analytical dipole model was developed. For this model, an accurate magnetic moment was obtained by characterising a 1 mm LCS ball bearing using a Superconducting Quantum Interference Device Vibrating Sample Magnetometer (SQUID-VSM) to get a value of $0.87039 \times 10^{-3} Am^2$. Using the dipole equation $b(r,m) = (\frac{u_0}{4\pi||r||^3}(3\hat{r}\hat{r}^T - \mathbb{I}_3))m$, the magnetic field strength surrounding the ball bearing was computed at each point within a 3D grid domain defined by $P_b = [x,y,z]$, where $x,y,z = 128$. The resulting values were discretised into bins and mapped to greyscale intensities. This data was visualised as dynamic tiles, which combined into a two-dimensional MRI image for each orthogonal plane in Fig. 2a. A single column and row were extracted at the centre and scaled by a factor of 4.258×10^7 to convert the values from Tesla to Hertz, as per the experimental data, for field map plotting as in Fig. 2b.

RESULTS AND DISCUSSION

The experimental data presented in Fig. 1a illustrate the baseline signal profiles of a resin crosshair and a resin ball bearing, together with the corresponding signal voids generated by each structure. The resin of the crosshair is visible as it has fewer unpaired precessing protons than the agarose medium, creating a non-distorted signal void.

Comparing this to Fig. 1b shows the signal void for the 1 mm LCS ball bearing measured at approximately 14 mm, making the signal void approximately 13 times larger than the object itself. The ferromagnetic ball bearing was placed within a central cavity; tracing the crosshair axes to their intersection in Fig. 1d confirms that the ball bearing is visually centred at 12.5 mm from the crosshair edge. This is further verified in the custom Python 3D reconstruction of the noise in Fig. 1e and Fig. 1f, with 1e matching 1d. The resulting flower-shaped structure visible in Fig. 1f corresponds to the shape in Fig. 2a.

The analytical model was compared to the experimental data as in Fig. 2. The phase maps confirmed numerical agreement with the analytical model by matching the centre of the analytical model to that of the experimental data. Since the analytical model defines the ball bearing at its centre, this alignment confirms that the ball bearing is visually centred and quantitatively located at the centre of the signal void in the experimental dataset.

The algorithm was successfully evaluated by positioning the ball bearing at multiple locations within the crosshair, with an example shown in Fig. 2. The overlapped field maps gave a rounded average Root Mean Squared (RMS) Error between the analytical model and experimental data of 15.8%. This error is attributable to the additional signal noise produced from micro-bubbles, which is counteracted through the Gaussian smoothing.

Due to the magnetic similarity between LCS and iron, using a robot sleeve within the imaging domain from [2], the centering algorithm was successfully applied to get tip-position detection, as shown in Fig. 2d.

CONCLUSIONS

The analytical model represents the magnetic field generated by a centrally located 1 mm ball bearing and exhibits a low RMS error when compared with experimental data, confirming both numerically and visually that the ball bearing is positioned at the centre of the signal void within MRI images. At $\approx 2.44s$ a scan and an average of $0.072652s$ for the centering algorithm, the robotic application shows the potential of the proposed centering algorithm, alongside the MRI signal voids, to enable accurate localisation and future navigation of a continuum robot.

REFERENCES

[1] V. Lalande, F. P. Gosselin, M. Vonthron, B. Conan, C. Tremblay, G. Beaudoin, G. Soulez, and S. Martel, "In vivo demonstration of magnetic guidewire steerability in a mri system with additional gradient coils," *Med. Phys.*, vol. 42, no. 2, pp. 969–976, 2015.

[2] P. Lloyd *et al.*, "A concentric tube catheter for endoluminal interventions, steered and imaged via magnetic resonance imaging." *Comms Eng*, vol. 5, no. 74, 2026.

[3] M. E. Tiryaki and M. Sitti, "Magnetic resonance imaging-based tracking and navigation of submillimeter-scale wireless magnetic robots," *Adv. Intel. Sys.*, vol. 4, no. 4, p. 2100178, 2022.

[4] K. Zhang, A. J. Krafft, R. Umathum, F. Maier, W. Semmler, and M. Bock, "Real-time mr navigation and localization of an intravascular catheter with ferromagnetic components," *Magn. Reson. Mater. Phy.*, vol. 23, p. 153–163, 2010.

[5] R. Zhai, Z. Pan, Y. Kou, C. Yang, Y. Ruan, C. Xu, L. He, and J. Zang, "Multi-frequency dual-echo magnetic resonance imaging for real-time and artifact-free magnetic robot navigation," *Eng.*, vol. 57, pp. 189–199, 2026.

2 Step Constraint Resolution With Regularization For Robotic Endoscopes With Environment Contacts

M. Bianchi, B. Rosa, and H. Courtecuisse

ICube, UMR 7357 CNRS-Universite de Strasbourg, France

INTRODUCTION

Endoscopic ultrasound-guided gallbladder drainage (EUS-GBD) has become an increasingly popular option for managing acute cholecystitis (inflammation of the gallbladder), offering an alternative to surgical approaches, such as percutaneous gallbladder drainage or laparoscopic cholecystectomy. Following the French Society of Digestive Endoscopy (SFED) recommendations [1], EUS-GBD is rapidly expanding as it becomes one of the few possibilities in critical situations. The endoscope is inserted through a natural route, via the mouth, esophagus, stomach, and reaches the duodenum.

Robotic endoscopes are a kind of continuum robots, which have theoretically infinite Degrees of Freedom (DOFs) featuring an elongated structure that can be bent in a continuous curved shape allowing flexible access through small incisions or natural orifices. They are tendon-actuated robots, where tendons pull on a flexible backbone to create bending motion at the distal extremity. Those tendons are attached to the distal end and pass through holes in discs fixed to the backbone which allow for the tendon actuation to produce a curve on the backbone. The structure of the endoscope, while allowing for the navigation to deep areas with difficult access, present major challenges when simulated.

To accurately model the endoscope within an organism, constraints along the whole structure with external contacts as well as the friction between the tendons and the backbone needs to be taken into account. Those constraints introduce non-linearity which could over-constrain the system, rendering the problem ill-conditioned. Some existing works avoid this issue by assuming negligible friction between the tendons and the discs but literature shows that it induces effects like hysteresis that can't be neglected. To avoid the assumption of zero friction, the Finite Element Methods (FEM) can be used to represent our tendons and backbones. It provides an accurate contact model encompassing friction with the environment even though it might result in a complex system. In order to account for both external and internal contacts, usual solutions build one single constraint system. This means combining the complexities of both contact models. External forces lead to big systems due to the number of elements in FEM models while internal constraints are computed with small, strongly non linear systems. Combining both result in a big system with poor conditioning which results in a complex and time consuming system to solve.

In this paper, we explore a 2-step method to decouple the internal and external constraints resolution. At first, internal constraints are solved alone, with a Karush–Kuhn–Tucker (KKT) kind of resolution involving the complex tendons' behavior. The resulting Compliance Matrix is regularized to build a corrected system's solution. Then, external constraints are applied on this corrected solution, following a second KKT resolution. We show that applying regularization to the Endoscope system can reduce conditioning issues arising from internal constraints, thus mitigating the risk of propagating them to the final system.

MATERIALS AND METHODS

The endoscope used in this study is modelled by two pairs of antagonist tendons and a central backbone. The tendons form two orthogonal planes and the endoscope can rotate around the backbone's axis, allowing bending motion in all directions. The tendons are rigidly attached to the distal end of the backbone and slide through the discs otherwise. (See Fig.1)

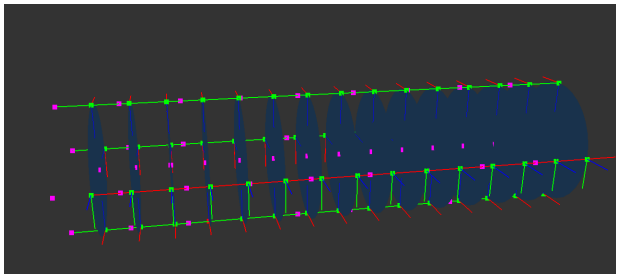


Fig. 1: Model used for the experiments with central backbone and 4 tendons. Physical Degrees of Freedom in purple, discs attachment points in green.

Both the tendons and the backbone are discretized using a finite-element formulation based on Timoshenko beam theory, where each beam is represented by a sequence of nodes connected by deformable beam elements. Each node has 6 DOFs: 3 translations and 3 rotations. Element stiffness matrices are computed from the beam's material and geometric properties, including Young's modulus, Poisson's ratio, density/mass, segment length, and cross-sectional radius. These parameters define the axial, bending, torsional, and shear behavior of the structure.

The unconstrained mechanical equilibrium is formulated as $Ax = b$, where x denotes the nodal displacement vector, b gathers the external loads and generalized forces

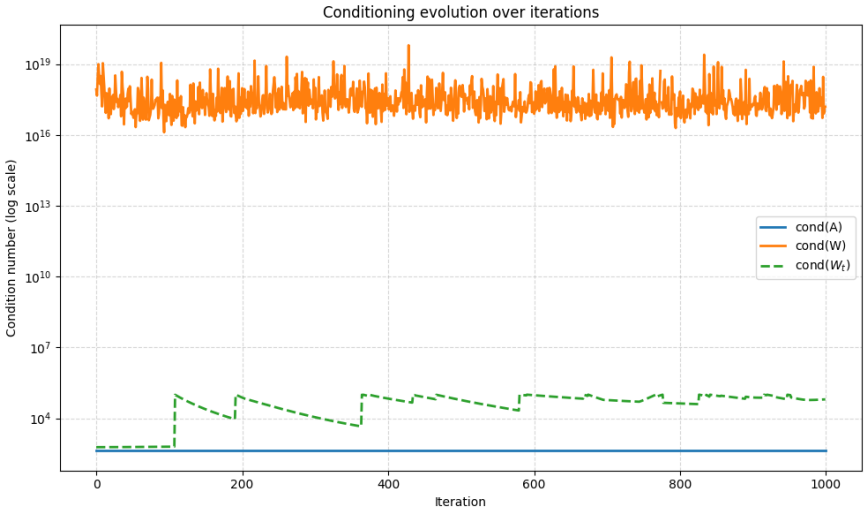


Fig. 2: Evolution of the conditioning of W, the regularized matrix W_t and A over 1000 iterations

applied to the system, and A represents the assembled global system matrix derived from the FEM discretization and time integration scheme [2][3]. The system is also subject to the constraint $Hx = \delta$. With $H = \frac{\partial C}{\partial x}$ the Constraint Jacobian and δ the constraint violation. Using Lagrangian multipliers the problem is formulated as :

$$\mathcal{L}(x, \lambda) = \frac{1}{2} x^T A x - b^T x + \lambda^T (Hx - \delta).$$

Applying the Karush–Kuhn–Tucker (KKT) conditions results in the following system

$$\begin{cases} Ax + H^T \lambda = b, \\ Hx = \delta. \end{cases} \tag{1}$$

H is constructed geometrically by projecting each backbone attachment point onto the corresponding tendon beam segment. A given backbone disc n is only intersected by one beam segment of a given tendon i. Let $p_{i,j}$ and $p_{i,j+1}$ denote the two extremities of this beam segment. The projected attachment point is interpolated as $p_c = \alpha p_{i,j} + \beta p_{i,j+1}, \alpha + \beta = 1$.

A local orthonormal frame $(\vec{t}, \vec{n}_1, \vec{n}_2)$ is constructed at p_c with $\vec{t}$ the tangent direction of the tendon segment and $(\vec{n}_1, \vec{n}_2)$ the plane orthogonal to $\vec{t}$. Tangential friction is omitted, the resulting constraints are :

$$n_k^T \left(x_n^{free} - \alpha x_{i,j}^{free} - \beta x_{i,j+1}^{free} \right) = 0, \qquad k \in \{1, 2\},$$

with x_n^{free}, $x_{i,j}^{free}$ and $x_{i,j+1}^{free}$ the positions of the backbone node n and the tendon nodes after unconstrained motion.

Solving the KKT system (1) yields

$$x = A^{-1} b - A^{-1} H^T \lambda, \qquad W\lambda \quad = HA^{-1} b - \delta$$

Solving this system requires the computation of the compliance matrix W^{-1}, with $W = H^T A^{-1} H$. Since H is non-linear, the matrix W may become ill-conditioned or rank-deficient. To mitigate this issue, we apply a SVD regularization by thresholding small singular values [4] to build W_t and solve $\lambda = W_t^{-1}(HA^{-1}b - \delta)$.

CONCLUSIONS AND DISCUSSION

We compare the conditionings W with and without regularization and we also compare with the conditioning of A to distinguish between numerical issues arising from the mechanical discretization and those induced by constraint redundancy.

The results in Fig.2 show that before regularization W has a very high conditioning which leads to major instabilities and errors since it isn't easily invertible. The regularization reduces the order of the conditioning drastically which is promising. This remains a preliminary work that needs to be compared to real data to ensure the proper modeling of the endoscope behaviour before integrating it into a constrained environments with organs and body motion.

The main idea is that by decomposing the constraint calculations, we reduce the propagation of numerical instabilities in the solution. Even though the internal constraint matrix may be poorly conditioned, its small size makes it easier to regularize at this stage before assembling the much larger finite element system of the environment. The final goal of this study is to present a model that combines high precision over micro behaviour of all components, accuracy of external interactions while preserving numerical stability and execution speed.

REFERENCES

[1] V. et al., ''Recommandations de pratique pour le drainage biliaire endoscopique en urgence,'' *Hépato-Gastro & Oncologie Digestive*, vol. 28, no. 10, pp. 1209–1218, 2021.

[2] H. Courtecuisse, ''Real-time Finite Element simulations for surgical assistance - applications to training, augmented reality and robotic control,'' Habilitation à diriger des recherches, Université de strasbourg, Oct. 2021.

[3] J. Przemieniecki, *Theory of Matrix Structural Analysis*, ser. Dover Civil and Mechanical Engineering, 1985.

[4] P. Hansen, ''Rank-deficient and discrete ill-posed problems,'' Ph.D. dissertation, 1996.

[5] C. Duriez, F. Dubois, A. Kheddar, and C. Andriot, ''Realistic haptic rendering of interacting deformable objects in virtual environments,'' 2008. [Online]. Available: https://arxiv.org/abs/0804.0561

Deep Learning 3D Surgical Tool Pose Estimation using Geometry-Aware Correction

A. Dimitriadis[1], G. P. Moustris[2], and C. S. Tzafestas[1,2]

[1]*National Technical University of Athens, Greece*
[2]*HERON - Hellenic Robotics Center of Excellence, Athena Research Center, Greece*

INTRODUCTION

In minimally invasive surgery (MIS), where visual feedback is provided through an endoscope, robust pose estimation of surgical instruments is a critical prerequisite for workflow analysis, intraoperative guidance, context-aware assistance, and automated skill assessment [1]. While deep learning models have significantly advanced 2D instrument localization, extending these capabilities to robust 3D tracking via stereo endoscopy remains a challenging.

Recent 2D pose architectures, such as unified multi-head networks and YOLO-based detectors, excel at extracting keypoints within individual frames [2]. However, treating 3D reconstruction as a passive downstream calculation based on these 2D outputs is inherently fragile. Zero-shot learning architectures have been utilized to generalize pose estimation to unseen surgical instruments, but their practical deployment is hindered by their requirement for 3D CAD models of the tools and reliable depth sensing. [3].

Building upon our 2D tracking framework presented in [4], which integrates YOLOv8m and RTMPose-m with an adaptive temporal filter, this work introduces a novel, active 3D reconstruction backend. To prevent depth (Z-axis) explosions prior to stereo triangulation, our backend deploys self-healing mechanisms, namely cross-camera reprojection and kinematic extrapolation, to mathematically correct transient visual failures, providing an uninterrupted 3D data stream for downstream robotic control.

MATERIALS AND METHODS

The surgical tool tracking framework (Fig. 1) is implemented as a four-stage hierarchical pipeline.

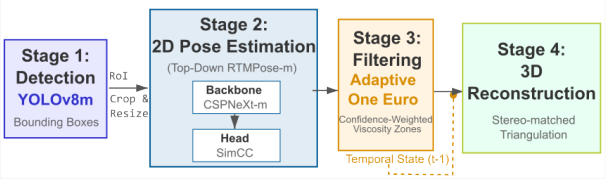


Fig. 1 The proposed framework: Tool detection, Pose estimation, temporal filtering and 3D reconstruction.

This project is funded by the European Union under Horizon Europe grant No. 101136568 - HERON

The models were trained and evaluated using the SurgPose dataset (https://github.com/zijianwu1231/SurgPose), a comprehensive benchmark for 2D and 3D surgical tool pose estimation, featuring a wide variety of robotic instruments. The dataset provides 2D keypoint annotations for 14 distinct keypoints (7 for each of the 2 manipulators), alongside tightly synchronized stereo-kinematic data, captured at 30 FPS with a spatial resolution of 1440×986 pixels.

A. Stage 1: Tool Detection via YOLOv8

We employ an anchor-free YOLOv8m model for initial tool localization, cropping the image to a localized region of interest (RoI) so the subsequent pose estimator focuses exclusively on the instrument. To ensure the tool edges are not truncated by overly restrictive bounding boxes, we apply a safety padding of $\delta = 40$ pixels to every bounding box.

B. Stage 2: Pose Estimation via RTMPose

The localized RoI is processed by the RTMPose-m model to identify 14 keypoints defining the tool's structure. We utilize the CSPNeXt-m backbone, enhanced with channel attention modules to better distinguish instrument features from the complex tissue background.

C. Stage 3: Confidence-Weighted Temporal Filtering

To enhance stability during high-frequency tracking, we implement a post-processing stage using an adaptive *One Euro Filter* on the 2D keypoints extracted from both the left and right stereo optical frames. This temporal filtering is an essential for robust stereo reconstruction due to the nonlinear nature of 3D error propagation. Because depth estimation relies heavily on the horizontal pixel disparity between stereo pairs, even minor sub-pixel 2D jitter in a single camera view is dramatically amplified during triangulation.

D. Stage 4: Geometrically Constrained 3D Reconstruction

To reconstruct the 3D spatial trajectory, triangulated camera coordinates are mapped to the robot base frame using a trial-specific rigid transformation matrix derived from ground truth. To prevent depth (Z-axis) explosions and subsequent history corruption inherent to naive stereo triangulation, our 3D backend employs three active mechanisms:

- **Pre-Triangulation Anomaly Detection**: Continuously monitors the 2D keypoint stream for spatial anomalies—such as sudden pixel jumps, neural

network confidence drops, or physically impossible stereo disparities—suspending naive triangulation before mathematical errors can propagate.

- **Cross-Camera Reprojection**: Handles asymmetric tracking failures by reconstructing a compromised 2D keypoint from its valid contralateral observation. For example, if the left tracking stream fails at time t while the corresponding right image keypoint $\mathbf{x}_{r,t}$ remains reliable, the missing left keypoint $\hat{\mathbf{x}}_{l,t}$ is synthesized by back-projecting $\mathbf{x}_{r,t}$ using the most recent valid depth anchor Z_{t-1} and then forward-projecting the resulting 3D estimate into the left camera:

$$\hat{\mathbf{x}}_{l,t} = \pi_l \left(\pi_r^{-1} \left(\mathbf{x}_{r,t}, Z_{t-1}; \mathbf{P}_r \right); \mathbf{P}_l \right). \quad (1)$$

  Here, $\mathbf{P}_r$ and $\mathbf{P}_l$ denote the right and left camera projection matrices, respectively.
- **Thresholded Kinematic Extrapolation**: Compensates for a total visual collapse when both cameras are blinded simultaneously ($c_l, c_r < \delta_{\text{conf}}$). The system ignores corrupted image data and advances the tool's 3D position linearly using a smooth velocity vector $\mathbf{v}_t$:

$$\begin{aligned} \mathbf{P}_{\text{extrap}} &= \mathbf{P}_{\text{cam},t-1} + k \cdot \mathbf{v}_t, \\ \mathbf{v}_t &= 0.2\,\Delta\mathbf{P}_{t-1} + 0.8\,\mathbf{v}_{t-1} \end{aligned} \quad (2)$$

  This extrapolated 3D coordinate is subsequently forward-projected to both image planes to sustain the tracking regions of interest until visual recovery.

RESULTS

The 2D keypoint estimation frontend was evaluated on the dedicated test set utilizing the standard Object Keypoint Similarity (OKS) metric. Consistent with the 2D training configuration, performance benchmarks for this stage were established exclusively using the left camera frames.

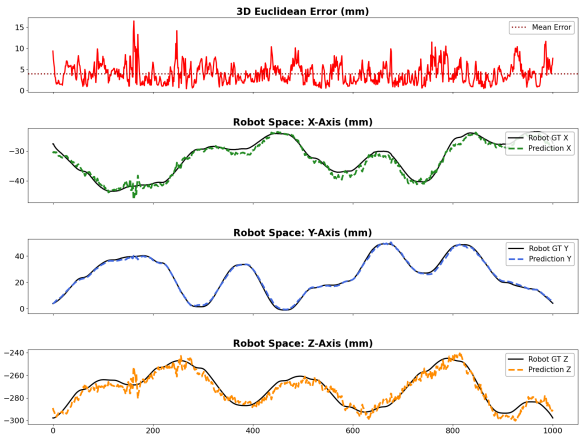


Fig. 2 3D tracking performance and coordinate decomposition (Trial 000026, PSM1)

We report the Average Precision (AP) using a per-keypoint constant $\sigma = 0.05$ for all 14 keypoints. The quantitative results are summarized in Table I.

Our model demonstrates strong performance by achieving high average precision across most instruments, validating the architecture's ability to extract robust features from distinct tool geometries. However, performance degradation is observed when encountering extreme variations in tool scale and structural morphology.

TABLE I Comparison of AP_{50-95} on SurgPose Dataset

Architecture	Ovr	LND	MND	M.For	Sci	For	Pro
HRNetV2p	64.9	**84.7**	45.9	55.6	66.0	48.8	88.3
MID-POSE [2]	71.4	81.2	**55.9**	68.5	**95.0**	56.0	72.1
Ours	**73.0**	73.8	15.8	**79.6**	71.5	**97.2**	**99.9**

TABLE II Global 3D tracking error and Average Distance of Model Points (ADD) accuracy over 14 trials.

Manipulator	Mean 3D ± SD (mm)	Mean X (mm)	Mean Y (mm)	Mean Z (mm)
PSM1	12.64 ± 45.16	2.48	2.12	11.80
PSM3	9.29 ± 8.94	1.38	2.56	8.48

Manipulator	ADD @1.0 mm	ADD @2.5 mm	ADD @5.0 mm
PSM1	2.06%	18.97%	45.15%
PSM3	1.54%	15.68%	40.16%

By successfully suppressing erratic frame-to-frame displacements, the filtering strategy minimizes jitter during challenging tracking conditions to produce a smooth 3D kinematic profile that closely mirrors the natural, intended instrument trajectories captured in the ground truth.

As demonstrated in Table II, the overall tracking discrepancy in 3D Reconstruction is heavily dominated by errors along the optical axis; the Mean Z error accounts for the vast majority of the spatial deviation, whereas the tracking errors in the X and Y planes remain consistently low at under 2.6 mm.

DISCUSSION

This study presents a robust, four-stage pipeline extending surgical tool tracking from 2D keypoints to continuous 3D trajectories. Combining a deep learning visual frontend with an active, geometry-aware 3D backend effectively mitigates the depth (Z-axis) explosions inherent to naive stereo triangulation. While our self-healing heuristics yield stable coordinates, future work will transition these rule-based safeguards into data-driven ML models within the backend geometry engine to handle severe occlusions. Additionally, to resolve our current dependence on static, trial-specific rigid transformations, we aim to integrate an online hand-eye calibration engine, such as [5], to dynamically compensate for intraoperative mechanical drift.

REFERENCES

[1] W. Wei *et al.*, "Instrument detection and pose estimation in surgical interventions," in *Medical Imaging with Deep Learning*, 2025.

[2] ——, "MID-POSE: Multi-instrument detection and pose estimation in endoscopic surgery," 2026.

[3] U. Rai, H. Xu, and S. Giannarou, "Surgpose: Generalisable surgical instrument pose estimation using zero-shot learning and stereo vision," 2025. [Online]. Available: https://arxiv.org/abs/2505.11439

[4] A. Dimitriadis, G. P. Moustris, and C. S. Tzafestas, "2d surgical tool pose estimation via top-down deep learning and adaptive temporal filtering," in *18th Hamlyn Symposium on Medical Robotics*, 2026.

[5] Z. Wu, S. Yang, Y. C. Lee, E. Prisman, Y. Jin, and S. E. Salcudean, "Surgcalib: Gaussian splatting-based hand-eye calibration for robot-assisted minimally invasive surgery," 2026. [Online]. Available: https://arxiv.org/abs/2603.08983

Exploiting Kinematic Redundancy for Active Stiffness Modulation of Continuum Concentric Push-Pull Robots

Azad Artinian[1], M. Taha Chikhaoui[2], and Benoit Rosa[3]

[1] *Sorbonne Université, CNRS, INSERM, ISIR UMR7222, Paris, France.*
[2] *Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, 38000 Grenoble, France*
[3] *Université de Strasbourg, CNRS, INSERM, ICube UMR7357, Strasbourg, France.*

INTRODUCTION

Recent progresses in modeling and control of continuum robots (CRs) has enabled submillimeter accuracy in tip position control. Despite these advances, several challenges still hinder their widespread adoption. One key challenge is the ability to modulate their stiffness independently from their shape. Many applications require varying stiffness levels at different stages of operation. For example, in minimally invasive surgey (MIS) or endoscopic procedures, a compliant behavior is advantageous for safely navigating through the anatomy. However, increased stiffness becomes necessary during surgical manipulation or when inserting instruments, in order to prevent undesired deflections caused by external disturbances or interactions. While robots can be rendered more compliant through control strategies such as impedance control, CRs are not inherently stiff. The challenge, therefore, lies in increasing their stiffness when needed. So far in the litterature, stiffness increase of CRs have relied on additional embedded modules, such as sheaths [1] and locking mechanisms [2] that require extra actuators, significantly increase the CR's overall dimension, and typically regulate stiffness in an on/off fashion. Such approaches are therefore not easily compatible with MIS applications. The approach presented in this abstract leverages key properties of continuum concentric push-pull robots (CPPRs) to smoothly increase (up to two times experimentally) the stiffness of CPPRs in the main bending direction while avoiding significant tip motion. Notably, the presented approach solely relies on CPPR's own actuation, without design modifications or additional modules.

MATERIALS AND METHODS

A. CPPR Principle

CPPRs are an emerging type of continuum robots. The two-tube version studied in this abstract consists of two notched concentric tubes attached at their end. The notches modify the material distribution such that, when assembled, pushing or pulling the inner tube wrt. the outer tube creates a bending motion (Fig. 1 left and middle). Furthermore, applying a torque along the z-axis in Fig. 1 on the inner tube wrt. the outer tube creates an out-of-plane motion (Fig. 1 right). CPPRs therefore have 3 actuation variables, the force F_2 and torque C_2 applied on the inner tube, and the simultaneous rotation θ_{12} around the z-axis of both tubes. We define $\Gamma = [F_2, C_2, \theta_{12}]$ the configuration of a CPPR, and $r(\Gamma)$ the resulting tip position. In the presented approach, we use the CPPR model proposed in [3], based on the Cosserat rod theory. The proposed method is tested on the CPPRs designs $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ from [3] where their physical characteristics are detailed further. They are 3D-printed in polylactic acid (PLA) using a Raise3D Pro2 printer with a 0.4 mm nozzle.

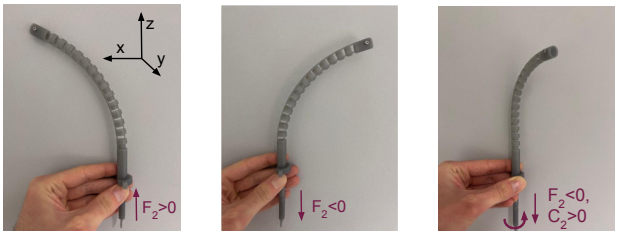
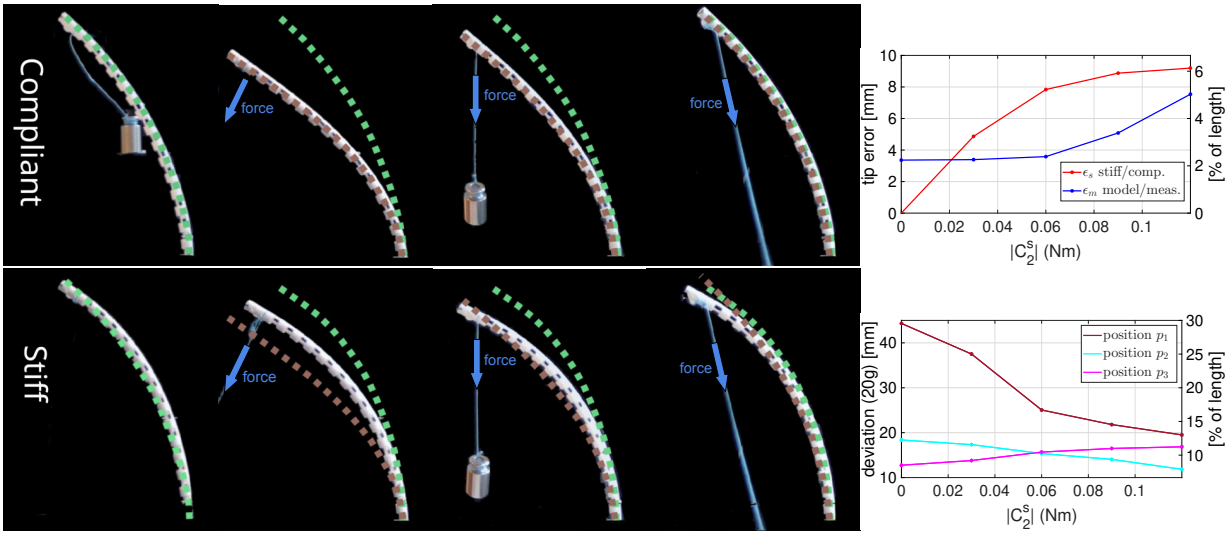


Fig. 1: Photo of the two-tube CPPR when pushing, pulling, and simultaneously pulling and rotating the inner tube with respect to the outer tube.

B. Method

Establishing an analytical relationship between the torque C_2 and the tip stiffness is challenging since CPPR models [3] do not have a closed-form analytical solution and no explicit analytical relation links the actuation variables to the strains. Our method is therefore built on two key observations. First, the CPPRs' workspace describes a 2D surface in 3D. Since CPPRs have 3 independent actuators, this makes them kinematically redundant. There are therefore infinite close configurations Γ^1, Γ^2, etc that produce similar tip positions, i.e. $r(\Gamma^1) \approx r(\Gamma^2) \approx ....$ The second observation is that applying a torque C_2 both creates an out-of-plane tip motion, and also increases the stiffness of the CPPR in the main bending direction. The reason is twofold: (i) Applying C_2 injects potential elastic energy that is stored in the system. (ii): It creates a torsion of the inner tube which modifies the material distribution due to the notches, thus increasing the stiffness in the main bending direction while decreasing it in the out-of-plane direction. By combining these two properties, the stiffness of the CPPR can be increased by applying a torque C_2 on the inner tube, while exploiting kinematic redundancy to adjust the other actuation variables, F_2 and θ_{12}, so as to compensate for the tip displacement.



(a) Unloaded (b) Weight in position p_1 (c) Weight in position p_2 (d) Weight in position p_3 (e) Error and deviation wrt. θ_2

Fig. 2: Deviation of a CPPR under a 20g weight ($\mathcal{R}_1$'s payload) in p_1, p_2, and p_3, for both compliant Γ^c and stiff Γ^s configurations. Green: unloaded compliant reference configuration, Brown: loaded compliant configuration.

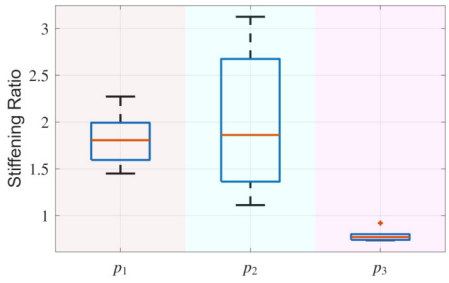


Fig. 3: Tip stiffness increase along each direction

RESULTS

To evaluate our approach, we define an initial compliant configuration Γ^c and a target stiff configuration $\Gamma^s = [F_2^s, C_2^s, \theta_{12}^s]$. The maximum torques C_2^s applied to stiffen the system are respectively 0.12, 0.8 and 0.6 Nm for $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$. We then solve the optimization problem in Eq. (1) to obtain F_2^s and θ_{12}^s that compensate for the shift in tip position induced by applying C_2^s.

$$(F_2^s, \theta_{12}^s) = \arg\min_{F,\theta} ||r([F_2, C_2]) - r([F, C_2^s, \theta])||, \quad \text{subject to : } |F_2^s| \leq F_2^{max} \tag{1}$$

The resulting Γ^s therefore increases the stiffness with only marginal changes to the CPPR's tip position. Both configurations are then applied to CPPR prototypes. To apply external tip forces f_+ from various directions, a weight is placed at one of three predefined locations p_1, p_2, or p_3: $f_+(p_1)$ and $f_+(p_2)$ lie in the CPPR's primary bending plane, while $f_+(p_3)$ is along the out-of-plane direction. The tip deviation under the weight for both compliant and stiff configurations are then compared as represented in Fig. 2. On 13 successive tests with different Γ^c on $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$, we obtained a mean stiffness increase ratio of 1.81, 2 and 0.79 for the weights in p_1, p_2, p_3 (see Fig. 3).

CONCLUSIONS AND DISCUSSION

In conclusion, we propose an MIS compatible method to increase the stiffness of CPPRs while preserving the tip position. It relies on injecting elastic potential energy into the system through torque actuation, combined with modifying the material distribution to increase stiffness. The resulting tip displacement is compensated by exploiting kinematic redundancy. This strategy achieves a twofold increase in stiffness along the primary bending direction, with the trade-off of a slight reduction in out-of-plane stiffness, which remains high in absolute terms. Using Nitinol tubes is a promising future direction, as they store more energy without breaking, further increasing stiffness. Simulation showed stiffening ratios up to 9×; overcoming the 3D-printed prototypes' material limits could help reach similar performance in practice. Furthermore, the optimization problem in Eq (1) is currently solved offline. A closed-loop, online control strategy would be a valuable extension, allowing stiffness to be automatically tuned for optimal performance during the surgical procedure.

REFERENCES

[1] M. Langer, E. Amanov, and J. Burgner-Kahrs, "Stiffening sheaths for continuum robots," *Soft Robotics*, vol. 5, no. 3, pp. 291–303, 2018.

[2] C. Yang, S. Geng, I. Walker, D. T. Branson, J. Liu, J. S. Dai, and R. Kang, "Geometric constraint-based modeling and analysis of a novel continuum robot with shape memory alloy initiated variable stiffness," *The Int. J. Robot. Res.*, vol. 39, no. 14, pp. 1620–1634, 2020.

[3] M. Tummers, F. Boyer, V. Lebastard, A. Offermann, J. Troccaz, B. Rosa, and M. T. Chikhaoui, "Continuum concentric push–pull robots: A cosserat rod model," *The Int. J. Robot. Res.*, vol. 44, no. 2, pp. 216–246, 2025.

Acknowledgments: This work is supported by grants ANR-18-CE19-0012, ANR-11-LABX-0004-01, ANR-21-ESRE-0015 and ANR-10-IAHU-0002.

A SOFA-Based Virtual Reality Simulator for Vascular Procedures Integrated with the da Vinci Research Kit

G. Gamberini[1,2], C. Lodin[2], S. Tognarelli[2], A. Menciassi[1,2]

[1]*Health Science Interdisciplinary Centre, Scuola Superiore Sant'Anna*

[2]*The BioRobotics Institute, Scuola Superiore Sant'Anna*

giulia.gamberini@santannapisa.it

INTRODUCTION

Digital twins (DTs) are virtual replicas of physical systems, reproducing their real-time behaviour and responses. In healthcare, DTs support personalised medicine, surgical planning, and training, particularly in Robotic Assisted-Surgery, by enabling risk-free practice in virtual environments [1], [2].

In surgical training, DTs can be virtual replicas of patient anatomy or physical simulators, combining the reproduction of mechanical responses, collisions and interaction dynamics [3]. However, their development is challenged by the computational complexity of contact-rich manipulation, involving interactions between rigid and soft bodies [3].

Existing VR simulators, such as those emulating the da Vinci Surgical System (Intuitive Surgical Inc., USA) or the da Vinci Research Kit (dVRK), use platforms like Coppelia Robotics, Unity, or PhysiX5 [4], [5]. Commercial solutions, like the da Vinci Skills Simulator or Simbionix LAP Mentor, offer standardised training modules but lack customisability and real-time adaptability to patient-specific anatomy. Recent works have also introduced digital twin simulators with both physical and virtual counterparts, such as a needle training task board [6] and a VR-based digital-twin robotic minimally invasive surgery simulator [7]. Simulation environments enable physically based simulations for tasks such as suturing and laparoscopy [8].

This work presents the realization of a virtual replica of a high-fidelity vascular physical simulator [9], built on Simulation Open Framework Architecture (SOFA) [10], with ad-hoc algorithms for the cutting task and integrated with the dVRK. The final objective is creating a digital twin simulator to augment the capabilities of the physical one and to realize personalized training curricula by exploring alternative scenarios.

MATERIALS AND METHODS

The simulation scene was built using SOFA, and the simulated objects were modelled using Fusion 360 (Autodesk, USA). The virtual simulator was designed as a digital replica of the high-fidelity physical simulator developed in [11], which features a sensorized silicone vessel replicating a pulmonary vein fixed between two anchoring systems and covered with an adipose surrounding tissue. The procedure for which the system was realized involves removing the simulated adipose tissue, creating a passage beneath the vein, and performing the resection of the vessel—mirroring clinical tasks like cancerous or damaged tissue removal [11]. This approach ensures consistency between virtual and physical training experiences. The virtual simulator faithfully replicates the physical simulator's objects and tasks while employing optimised meshes to ensure real-time performance (>20 Frame Per Second (FPS)). It operates on a portable laptop (Intel i7-10750H, 16GB RAM, Ubuntu 20.04.1) without requiring a CUDA-compatible GPU. SOFA v23.06.00 was selected for its robust real-time soft-body simulation capabilities using the Finite Element Method (FEM). SOFA's Python3 integration (SofaPython3) enabled the implementation of custom logic and dynamic scene modifications. A Robot Operating System (ROS) bridge was developed to facilitate bidirectional communication between the dVRK and the simulation environment. SOFA's hierarchical scene modularly organises mechanical, visual, and collision models for each object which coexists via mappings, thus ensuring synchronised behaviour. This architecture supports modular assembly of solvers and components. In our simulated scene, the vein features an elliptic cylindrical shape, and it is implemented using a `MechanicalObject`. During the simulated surgical task, the adipose tissue needs to be both cut and grasped. ***Cutting*** (i.e., *grasping* and *carving*) behaviour is not natively supported in SOFA; thus ad-hoc algorithms were developed using the `CarvingManager` (`CM`) component. In addition, a topology modifier component was added to the adipose tissue to enable the `CM` to modify the object topology during the simulation. *Grasping* is implemented using a custom controller that creates constraints between the gripper and tissue degrees of freedom (DOFs) upon collision, temporarily disabling physical collisions to avoid instabilities, while *carving* is triggered by the "+" dVRK pedal, removing tetrahedrons from the adipose tissue mesh. The virtual simulator integrates with the dVRK via ROS, subscribing to topics for gripper pose, jaws angle, and pedal state, with custom callbacks preprocessing data (e.g., positional offsets, quaternion adjustments) to update the simulation state. Regarding the system architecture, the dVRK computer acts as the ROS master, while the simulation registers a node to exchange real-time data. To optimise performance, only the dVRK grippers (end effectors) were simulated, controlled via a `RestShapeSpringsForceField` that dynamically updates based on ROS data, ensuring

stability in FEM computations. Strain computation tracks DOFs along the vessel to estimate deformation, with rupture occurring at 20% strain [12], and a warning system (LEDs) alerts user when threshold is exceeded. Vessel breakage is simulated by deactivating the original model and activating two new linear approximations of the deformed vein segments. This approach balances realism and computational efficiency while replicating clinical tasks such as tissue manipulation and vessel resection. Two distinct simulated scenarios were implemented:

- In the first scenario the adipose tissue was partially pre-cut with a small window above the vessel, thus enabling the user to familiarize and to focus exclusively on the vessel-lifting task using the grasper.
- The second scenario represented the complete simulation setup, incorporating an intact adipose tissue mesh, in which the user was required to grasp and cut the adipose tissue prior to exposing and lifting the vein.

To preliminary evaluate the usability of the system and the user adaptability, tests with twelve non-clinical users were performed. All the participants signed an informed consent form prior to the testing session. The users were required to perform two repetitions of the simulated task without the adipose tissue (i.e., familiarization – first scenario) and three complete tasks (second scenario). Time, vessel breakage flag, and a questionnaire answers were collected during the testing session.

RESULTS

In the final simulation design, the virtual simulator scene comprises the visual model of simulator, the vessel surrounded by the adipose tissue and the two grippers. Moreover, the simulator base hosts three LEDs that are used to provide feedback to the user regarding the deformation level applied to the vessel (Figure 1).

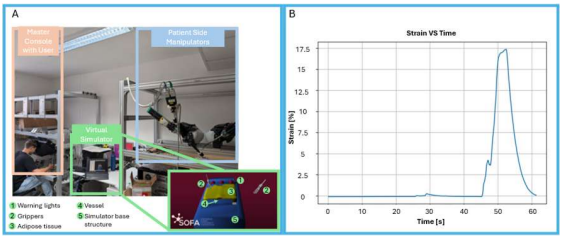


Figure. 1. (A) experimental setup used during the usability tests and VR simulator components; (B) example of strain over time plot recorded through the VR simulator.

The twelve users were all able to fully operate and control the simulator using the dVRK without experiencing any perceptible latency in all the 36 full repetitions, with an average refresh rate (mean ± standard deviation) of $(21.4 \pm 2.1)\ FPS$. The participants were engineering master and PhD students or researchers. Nine out of twelve: (*i*) declared the system is easy to use, (*ii*) declare that most people would learn to use the system quickly, (*iii*) felt confident in using the system. Moreover, all users increased their self-confidence through repetitions and in their experiences, they were feeling as they were controlling the robot. Eight people reduced the time needed to complete the task when comparing the first and the last repetitions. Overall, the mean completion time was 56.75 s during the first repetition, 50.75 s during the second repetition, and 42.83 s during the third repetition. Moreover, comparing the completion times between the first and the last repetitions using a Wilcoxon signed-rank test yielded a p-value of 0.0259, indicating a statistically significant reduction in task completion time.
In addition, thanks to the LEDs warnings none of the users broke the vessel during the last full trial.

CONCLUSION AND DISCUSSION

The work presented resulted in a fully functional virtual simulator based on open source only software, controlled by the dVRK platform and closely replicating and augmenting the high-fidelity physical simulator counterpart. This gives the basis to the realization of a digital twin simulator in which information flows continuously between the real and virtual systems. Additional usability testing sessions will be organised to advance the validation of the system.

REFERENCES

[1] Mekki YM, et al. "Digital twins for the era of personalized surgery." npj Digit Med; 2025; 8(1):283.

[2] Asciak L, et al. "Digital twin assisted surgery, concept, opportunities, and challenges." npj Digit Med; 2025; 8(1):32.

[3] Ou Y, et al. "A Realistic Surgical Simulator for Non-Rigid and Contact-Rich Manipulation in Surgeries with the da Vinci Research Kit" in Proceedings of 2024 21st International Conference on Ubiquitous Robots (UR); 2024; 64-70.

[4] Ferro M, et al. "A Portable da Vinci Simulator in Virtual Reality" in Proceedings of IEEE International Conference on Robotic Computing (IRC); 2019; 447-448.

[5] Fan K, et al. "A Unity-based Da Vinci Robot Simulator for Surgical Training" in Proceedings of IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob); 2022; 1-6.

[6] Caccianiga G, et al. "An Evaluation of Inanimate and Virtual Reality Training for Psychomotor Skill Development in Robot-Assisted Minimally Invasive Surgery." IEEE Trans Med Robot Bionics; 2020; 2(2):118-129.

[7] Cai X, et al. "Implementation of a Virtual Reality Based Digital-Twin Robotic Minimally Invasive Surgery Simulator." Bioengineering; 2023; 10(11):1302.

[8] Jourdes F, et al. "Visual Haptic Feedback for Training of Robotic Suturing." Front Robot AI; 2022; 9:800232.

[9] Gamberini G, et al. "Design and preliminary validation of a high-fidelity vascular simulator for robot-assisted manipulation." Sci Rep; 2024; 14(1).

[10] Payan Y. "Soft Tissue Biomechanical Modeling for Computer Assisted Surgery." in Studies in Mechanobiology, Tissue Engineering and Biomaterials; 2012; 11.

[11] Cao C, et al. "Incidence, Management, and Outcomes of Intraoperative Catastrophes During Robotic Pulmonary Resection." Annals of Thoracic Surgery; 2019; 108(5).

[12] Camasão DB, Mantovani D. "The mechanical characterization of blood vessels and their substitutes in the continuous quest for physiological-relevant performances. A critical review." Materials Today Bio; 2021; 10:100106.

Diffusion-Driven Safe Path Generation for Dual-Arm Surgical Autonomy

Bita Azad[1] and Siobhan Oca[2]

[1]*Department of Computer Science, University of North Carolina at Chapel Hill, USA*
[2]*Mechanical Engineering & Materials Science Department, Duke University, USA*

INTRODUCTION

Surgical robotics is evolving from precise translation of surgeon movement towards increasing autonomy and performing tasks with greater independence. While many single-arm tasks (e.g. gauze retrieval) can be automated reliably, clinically realistic workflows require coordinated bimanual manipulation. Diffusion-based models are increasingly used to learn safe, smooth robotic motion from demonstrations [1]. Still, dual-arm interactions such as needle handovers or bimanual regrasping have stringent requirements on synchronization, alignment, collision avoidance, and smooth motion. This makes robust bimanual control challenging, yet essential, for safe autonomous robot-assisted surgery [2].

To address complexities of bi-manual manipulation, this work introduces a hierarchical diffusion-based framework that decomposes dual-arm surgical tasks into short-horizon subtasks to enable safe and smooth motion execution. The approach combines high-level subgoal planning with low-level motion refinement to generate candidate waypoint sequences that respect geometric constraints important for surgical safety; although trajectories are represented in Cartesian space, RCM constraints [3] are enforced at execution via inverse kinematics, and iteratively denoising refines them to produce smooth, collision-free motion. The method generates safe trajectories under surgical constraints, validated on bimanual tasks, with available code attached[1].

MATERIALS AND METHODS

In this work, we introduce a diffusion-driven framework for generating safe, smooth dual-arm surgical trajectories for collision-free needle handover in constrained environments. We formulate dual-arm surgical manipulation as conditional waypoint generation in customized simulated workspaces with task-specific obstacles representing surgical constraints as shown in Figure2. Given structured numerical context describing the environment and interaction type, the diffusion planner generates waypoint sequences for object transfer while maintaining collision avoidance, smoothness, and kinematic continuity. At inference, trajectories are sampled conditioned on this context, as shown in Figure 1.

Trajectories are evaluated in a constrained surgical simulation with obstacle cavities. Two scenarios are evaluated: cavity passage and obstacle evasion, requiring smooth, collision-free execution. Safety verification is performed at the waypoint level. For cavity scenarios, alignment constraints ensure that the needle passes through the opening with sufficient clearance. For evasion scenarios, a minimum distance constraint is enforced between the needle and obstacle surface throughout the trajectory. Each waypoint encodes the Cartesian position, pitch orientation, and gripper state of both patient-side manipulators (PSM1 and PSM2), forming a 10-dimensional representation, which follows the SurRoL [4] simulator interface and captures the motion required for the task. A trajectory is defined as a sequence of such waypoints. Waypoint density is fixed per task to ensure reliable collision detection while maintaining computational efficiency. Under the assumption that both arms grasp the needle at its tips, inter-arm collisions are unlikely; failures result from grasping failures or geometric collisions.

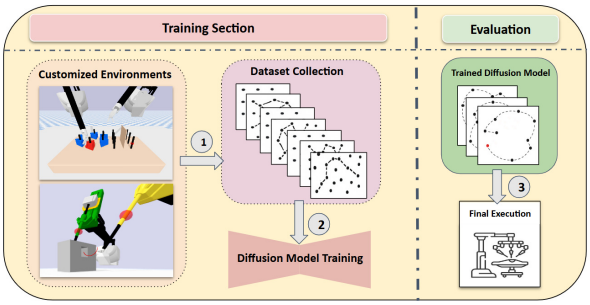
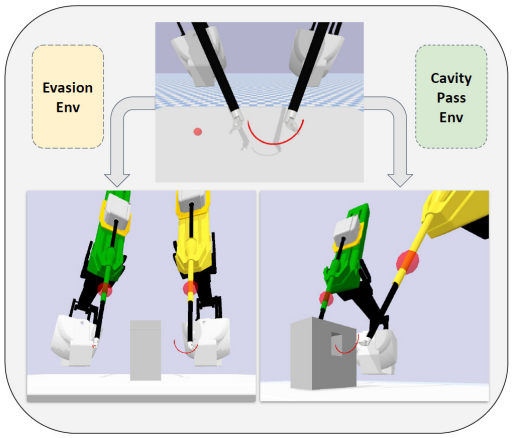


Fig. 1: Model Training and Inference Framework

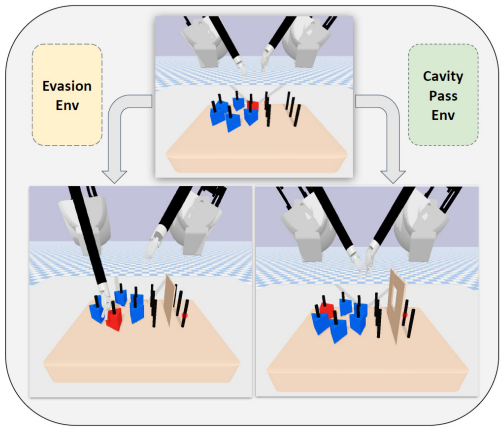
Trajectory demonstrations were generated in SurRoL, with each sample containing structured numerical observations, Cartesian waypoint actions, and task outcome labels. Data were collected across randomized obstacle cavity, obstacle evasion, and open-space transfer configurations, with testing performed on held-out configurations using 100 rollout trials per method and setting. We employ a denoising diffusion probabilistic model (DDPM) [5] to learn a conditional distribution over waypoint trajectories. During training, the model takes noised ground-truth waypoint trajectories together with task context and supervision signals (handover success, collision status, and smoothness) as input, and learns to predict the corresponding clean trajectories. Formally, the model learns the conditional distribution:

$$p_\theta(\tau_0 \mid \tau_t, \mathbf{c}), \tag{1}$$

[1]https://anonymous.4open.science/r/Safe-Path-Generator-76B5/README.md



(a) Needle re-grasp customized environment



(b) Bi-peg transfer customized environment

Fig. 2: Customized dual-arm surgical environments

where τ_0 denotes the clean waypoint trajectory, τ_t its noised version at diffusion step t, and $\mathbf{c}$ represents the conditioning context. At inference, the model starts from Gaussian noise $\tau_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and iteratively samples:

$$\tau_{t-1} \sim p_\theta(\tau_{t-1} \mid \tau_t, \mathbf{c}), \tag{2}$$

to generate new waypoint trajectories; optional guidance biases the sampling process toward safer and smoother motions without additional retraining.

CONCLUSIONS AND DISCUSSION

We evaluate our method in two obstacle-free and obstacle-present scenarios. In all experiments, each method executes waypoint trajectories through the same proportional controller in the simulator. Waypoints use a structured relative representation for learning and map to absolute end-effector states for execution. We compare our diffusion-based planner against an expert-designed waypoint controller (Ctrl), Behavior Cloning (BC) [6], and CHOMP [7], an optimization-based baseline for trajectory planning where smoothness and collision avoidance are enforced through cost-based optimization. RL baselines are not included, as this work focuses on generative waypoint planning for safe dual-arm coordination rather than reward-based control learning. In the results tables, WP denotes the number of waypoints, Fail and Succ indicate unsuccessful and successful episodes, Smth reports a normalized trajectory smoothness score across each task and its configuration, and Free, Evade, and Cav refer to collision-free, obstacle-evading, and cavity-constrained scenarios. The method outperforms all baselines across tasks. It achieves 100% success in free-space settings and improves success by 6–7% over CHOMP in constrained scenarios, with higher smoothness. Gains are strongest in cavity and obstacle avoidance cases, which shows the benefit of diffusion under constraints.

	Needle Regrasp				Bi-Peg Transfer			
Method	WP	Fail	Succ	Smth	WP	Fail	Succ	Smth
Ctrl (Free)	6	4	96	0.93	11	5	95	0.88
BC (Free)	6	4	96	0.95	11	4	96	0.91
CHOMP (Free)	6	2	98	0.96	11	3	98	0.93
Diff (Free) – Ours	6	0	**100**	**1.00**	11	0	**100**	**1.00**
Ctrl (Cav)	8	36	64	0.76	15	39	61	0.70
BC (Cav)	8	31	69	0.78	15	35	65	0.73
CHOMP (Cav)	8	29	71	0.79	15	34	66	0.76
Diff (Cav) – Ours	8	23	**77**	**0.83**	15	30	**70**	**0.82**
Ctrl (Evade)	11	33	67	0.55	18	34	66	0.43
BC (Evade)	11	30	70	0.59	18	31	69	0.50
CHOMP (Evade)	11	26	74	0.64	18	30	70	0.56
Diff (Evade) – Ours	11	19	**81**	**0.72**	18	27	**73**	**0.65**

TABLE I: Results on Needle Regrasp and Bi-Peg Transfer

In conclusion, we propose a diffusion-driven framework for safe path generation that conditions waypoint sampling on task success and trajectory smoothness to produce collision-free paths. This generative formulation captures the multimodality of successful behaviors and improves execution under tight geometric constraints, where deterministic baselines often commit to single or averaged trajectories.

REFERENCES

[1] B. Azad, S. Zargarzadeh, and F. Rudzicz, "Surgidiff: A context-aware diffusion recommender for safe surgical autonomy," in *2026 International Symposium on Medical Robotics (ISMR)*. IEEE, 2026, pp. 143–149.

[2] F. Zhong, Y. Wang, Z. Wang, and Y.-H. Liu, "Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2669–2676, 2019.

[3] N. Aghakhani, M. Geravand, N. Shahriari, M. Vendittelli, and G. Oriolo, "Task control with remote center of motion constraint for minimally invasive robotic surgery," in *2013 IEEE international conference on robotics and automation*. IEEE, 2013, pp. 5807–5812.

[4] J. Xu, B. Li, B. Lu, Y.-H. Liu, Q. Dou, and P.-A. Heng, "Surrol: An open-source reinforcement learning centered and dvrk compatible platform for surgical robot learning," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 1821–1828.

[5] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[6] F. Torabi, G. Warnell, and P. Stone, "Behavioral cloning from observation," *arXiv preprint arXiv:1805.01954*, 2018.

[7] M. Zucker, N. Ratliff, A. D. Dragan, M. Pivtoraiko, M. Klingensmith, C. M. Dellin, J. A. Bagnell, and S. S. Srinivasa, "Chomp: Covariant hamiltonian optimization for motion planning," *The International journal of robotics research*, vol. 32, no. 9-10, pp. 1164–1193, 2013.

[8] Z. Yang, C. Garrett, D. Fox, T. Lozano-Pérez, and L. P. Kaelbling, "Guiding long-horizon task and motion planning with vision language models," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 16 847–16 853.

Visual servoing of continuum robotic endoscope for laser-based path scanning

Xiao Xie[1], Gang Zheng[2], Isabelle Fournier[3], and Nageotte Florent[1]

[1]*ICube, UMR7357, CNRS, Université de Strasbourg, INSERM, France*
[2]*Defrost team, Inria Lille, CRIStAL UMR 9189*
[3]*PRISM, INSERM, U1192*

INTRODUCTION[1]

Accurate delineation of tumor margins is essential in minimally invasive surgery, where insufficient resection may lead to recurrence while excessive removal of healthy tissue degrades patient outcomes. New in vivo technologies such as mass spectrometry [1] enable real-time molecular tissue analysis but they rely on point-wise measurements, requiring precise scanning of tissue surfaces. Robotic endoscopes have been promising solutions to enable minimally invasive scanning in gastrointestinal applications [2]. However, precise manual teleoperation of continuum endoscopes for scanning tasks can be mentally demanding and time-consuming for the user, especially in confined anatomical environments.

In this work, our objective is to achieve autonomous non-contact tissue scanning using a continuum robotic endoscope coupled with the SpiderMass mass spectrometry system [1]. During scanning, an optical fiber used to locally vaporize tissues must be displaced along a path while maintaining a certain working distance to the tissue, with required accuracy of 1mm, to achieve effective tissue desorption. A guiding laser in the visible range and aligned with the desorption laser is used to guide the pointing. The scanning task is thus formulated as a dual-objective control problem involving path tracking in the image plane and regulation of the working distance between the laser source and the tissue surface. A hybrid visual servoing framework is proposed, where an estimation method allows accounting for modeling errors in the Jacobian.

MATERIALS AND METHODS

A. System Overview

The system consists of a continuum robotic endoscope equipped with a monocular camera and optical fibers, used for desorption and guidance, brought to the end of the endoscope through a channel. The position and orientation of the fibers at the distal side $(O, \mathbf{e})$ are fixed and calibrated beforehand offline from the observed laser positions at known distances (see Fig. 1). A laser triangulation method is used to estimate the distance d between the fiber tip and the tissue surface in real time, enabling depth perception at the measurement spot using the monocular camera. The endoscope has three actuated degrees of freedom (Dofs) at the distal side, including two orthogonal deflections and one translation (see Fig. 1). The deflections are obtained from the actuation of two pairs of antagonist cables, while the translation is obtained from the displacement of the endoscope body on the proximal side. The actuation variables at proximal side are denoted by $\mathbf{u}_p \in \mathbb{R}^3$, including two cable displacements and one proximal displacement. They are linked to the distal actuation variables $\mathbf{u}_d$ through a complex backlash/hysteresis function.

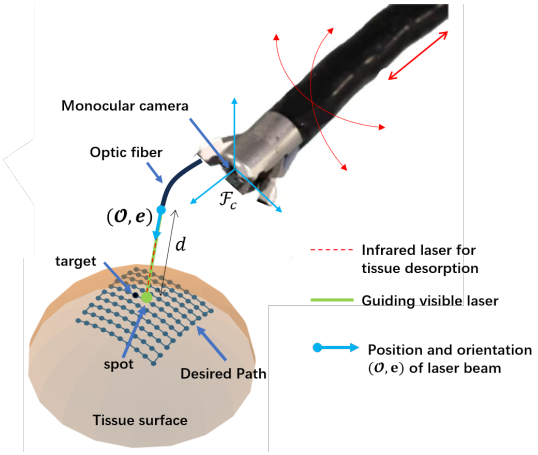


Fig. 1: Overview of the robotic system during a scanning task with a serpentine scanning path. Red arrows represent the three controlled Dofs at the distal tip.

B. Control Formulation

The control objective is to ensure that the laser spot follows a predefined path on the tissue surface while maintaining a desired working distance. The combined task error is expressed as a vector $\mathbf{y} = ((\mathbf{p}_s - \mathbf{p})^T, d - d_d)^T \in \mathbb{R}^3$, combining the error between spot position $\mathbf{p}_s$ and target position $\mathbf{p}$ in image frame and the error between distance d and desired working distance $d_d = 20$ mm. The desired distal control input is computed using inverse kinematics:

$$\mathbf{u}_d = J^{-1}\mathbf{y} \tag{1}$$

where $J \in \mathbb{R}^{3\times3}$ is the Jacobian matrix relating distal displacement $\mathbf{u}_d$ to the error variation $\mathbf{y}$.

C. Online Hybrid Jacobian Estimation

The Jacobian, given by $J = J^m + \delta J$, is decomposed into an analytical component and an unknown correction term which is estimated online (see Fig. 2). The analytical Jacobian J^m is derived from a piecewise constant curvature model (PCC) [3]. The residual term δJ captures structural inaccuracies in the model.

[1]This work was supported by INSERM through the program "Procédures chirurgicales et interventionnelles du futur"

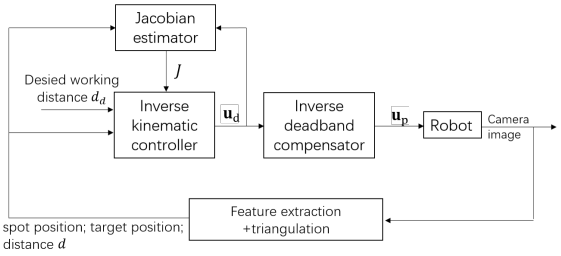


Fig. 2: Adaptive hybrid visual servoing control framework.

To estimate δJ, an adaptive square-root unscented Kalman filter (MASR-UKF) is employed. The state of the filter is defined as the vectorized Jacobian error $\mathbf{x} \in \mathbb{R}^9$, assuming that its evolution follows a random walk model

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{w}_k, \tag{2}$$

with $\mathbf{w}_k \in \mathbb{R}^9$ a Gaussian noise. This model reflects the quasi-static nature of the scanning task such that the Jacobian varies only slightly between successive time steps.

The observation model is constructed from the measurement of the task error:

$$\mathbf{z}_k = H_k((J_k^m)^{\mathrm{v}} + \mathbf{x}_k) + \mathbf{v}_k \tag{3}$$

where $\mathbf{z}_k = \mathbf{y}_k - \mathbf{y}_{k-1}$, H_k is the measurement matrix depending on the known control input $\mathbf{u}_d$ and $\mathbf{v}_k$ is a Gaussian noise.

A square-root formulation is adopted, in which the covariance matrix is propagated through its Cholesky factor, ensuring positive definiteness. Furthermore, an adaptive covariance estimation based on a modified Sage-Husa scheme is incorporated, which updates the process and measurement noise covariances online. As a result, the filter adjusts to time-varying uncertainties without requiring precise prior tuning.

D. Deadband Compensation

Cable-driven actuation introduces backlash, which affects control performance. Based on a simplified linear model, a deadband compensation is applied at change of motion direction: $u_p = u_d + \mathrm{sign}(u_d)\,\delta$ where δ is the estimated size of the deadband which captures the dominant lost motion during cable reversal and is identified offline [3]. The PCC model and the linear model of backlash provide real-time computational capability. In addition, the Jacobian estimation process is temporarily suspended during compensation to prevent estimation bias.

RESULTS

The performance of the proposed framework is evaluated through numerical simulations designed to replicate realistic surgical scanning conditions. The objective is to assess both path tracking accuracy and distance regulation under modeling uncertainties, sensing noise, and mechanical transmission effect (backlash). Kinematic uncertainties are introduced by perturbing the mapping between actuation inputs and robot configuration, resulting in 20% error in deflections which is consistent with the uncertainty measurement in [3]. Sensor noise is incorporated in both feature detection and laser spot localization, modeled as

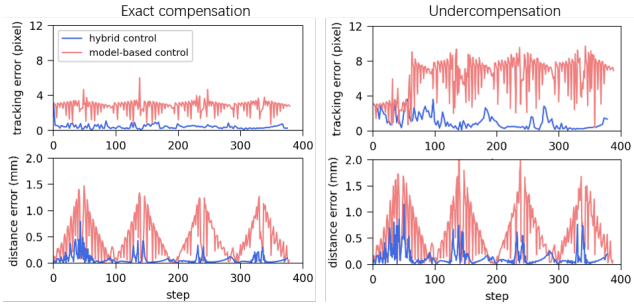


Fig. 3: Control error for path tracking and distance in case of exact or under-compensation of deadband. The serpentine path is predefined on a 4 cm × 4 cm region of a non-planar surface and illustrated in Fig. 1.

gaussian noises with 2 pixels of standard deviation. In addition, triangulation errors are introduced to reflect calibration inaccuracies and depth estimation limitations. Mechanical backlash is modeled as a deadband in the cable-actuation system. The identification error of deadband size δ is set to -25%, representing that the deadband is undercompensated.

The proposed hybrid controller is compared with a purely model-based approach. Performance is evaluated in terms of path tracking error and distance keeping error. The results show that, with the mean scan speed of 4 mm/s, the hybrid approach significantly improves accuracy (see Fig. 3). Taking the case of undercompensation as an example, the mean error in path tracking is 1.6 pixel in the image frame, equivalent to 0.66 mm in the euclidean space, providing a reduction of 82% compared to the model-based method. The mean error in distance keeping is 0.21 mm, representing a reduction of 87% compared to the model-based method.

These results validate the feasibility of the proposed framework in handling uncertainty and disturbances and highlight the importance of integrating estimation and compensation mechanisms in continuum robotic control.

CONCLUSIONS AND DISCUSSION

This work proposes a hybrid visual servoing framework for autonomous point by point tissue scanning using a continuum robotic endoscope equipped with a laser projection. The working distance is simultaneously controlled to allow correct acquisitions. Realistic simulations show that the hybrid Jacobian estimation significantly improves control accuracy compared to purely model-based approaches, allowing to reach accuracy compatible with medical use.

Future work will focus on experimental validation and integration with real-time tissue analysis systems.

REFERENCES

[1] P. Saudemont *et al.*, "Real-time molecular diagnosis of tumors using water-assisted laser desorption/ionization mass spectrometry technology," *Cancer cell*, vol. 34, no. 5, pp. 840–851, 2018.

[2] A. A. Nazari *et al.*, "Visual servoing of continuum robots: Methods, challenges, and prospects," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 18, no. 3, p. e2384, 2022.

[3] L. Ott *et al.*, "Robotic assistance to flexible endoscopy by physiological-motion tracking," *IEEE Transactions on Robotics*, vol. 27, no. 2, pp. 346–359, 2011.

Magnetic Actuation of Vine Robots for Tortuous Surgical Navigations

J. Davy[1], N. J. Greenidge[1], N. G. Kim[2], T. P. Dean[1], S. Kim[3], S. Park[2], B. Calme[1], P. Lloyd[1], J. H. Chandler[1], J. H. Ryu[2], T. K. Morimoto[3], and P. Valdastri[1]

[1]*STORM Lab, School of Electronic and Electrical Engineering, University of Leeds, UK*
[2]*Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea*
[3]*Department of Mechanical and Aerospace Engineering, University of California San Diego (UCSD), USA*

INTRODUCTION

Any decision to perform an endoscopic enabled intervention must be weighed against the risk of procedural complications. Standard endoscopic instruments are formed of hard bodies (relative to the anatomy through which they will travel) and thus carry the potential for traumatic tissue interaction. These instruments are typically controlled by internal tendons allowing limited bending focused purely at the tip and their insertion through the anatomy can be limited by frictional forces with surrounding tissues. Such tools are often characterised by low dexterity, limited control and large diameter.

These challenges motivate the design of new robotic instruments one of which is the everting catheter [1], [2]. These devices utilise an unravelling material which, instead of translating through a lumen, locomote via tip centred growth. This form of locomotion minimises potentially traumatic tissue interaction and focuses the forces of insertion at the distal end of the device [1]. Recently, such robots integrated with robotic steering methods such as pneumatic artificial muscles, have been designated as "vine robots" (so named due to their visual and behavioural similarity to the climbing plant) and have been explored for a wide range of applications.

In the medical domain, these vine robots form part of a wider group of soft robotic systems where it is hoped that softer materials can minimise stress on surrounding tissue during physical interaction. For target applications in the navigation of tight, tortuous lumens, a longstanding goal has been the miniaturization of such devices.

Magnetic actuation has emerged as an effective way of actuating miniaturised robot designs. This forms a wireless actuation methodology where the induced torques and forces are proportional to both the volume of internal magnetic material and the magnitude of the applied magnetic field and field gradients.

Over the past three years, we have developed several innovative approaches to magnetically-actuated vine robots. These unconventional robots are characterized by their minimal tissue interaction, the ability to generate forward propulsion, and precise steering achieved through controlled magnetic fields. This abstract provides an overview of three distinct approaches to the design of magnetically-actuated vine robots, each tailored to specific surgical navigation challenges. These distinct designs are compared in terms of their advantages and limitations, highlighting their suitability for various clinical applications paving the way for broader adoption of these robots in medical practice.

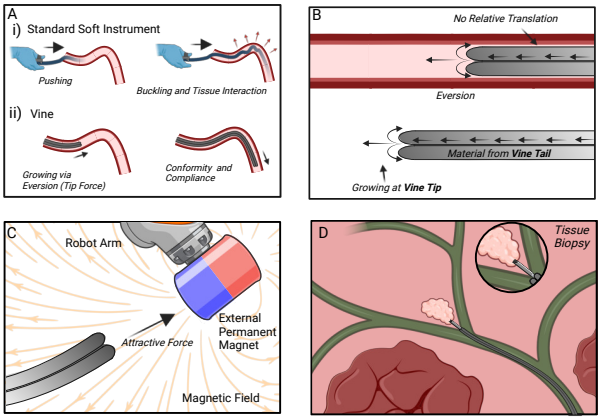


Fig. 1: An overview of magnetic vine robots.

I. APPROACHES

A. Hard Magnetic Tip: Kim et al. [3]

Our initial approach addressed the complex challenge of integrating a small permanent magnet at the tip of a vine robot. This task is complicated due to the everting nature of the vine robot, which makes it difficult to maintain a consistent tip location, as sections of the robot are continuously everted [4]. In collaboration with Jee Hwan Ryu's research group at KAIST, we successfully designed and demonstrated the first magnetic actuation system for vine robots, utilizing a magnetic tip (see Figure 2). This robot, with a diameter of 25 mm, features an endoscopic camera at the tip and includes a tool channel, irrigation channel, and an LED light source. We tackled the issue of the vine robot's fixed tip by pressurizing the vine channel internally, ensuring synchronization between the tether-attached robot tip and the vine's body.

A significant limitation of current magnetically actuated capsule robots is their insufficient forward propulsion force to overcome the accumulation of drag from the device tether. The vine robot's structure addresses this issue by applying a pneumatic tip force to propel the robot forward, reducing tether drag and adverse interactions with tissue due to the absence of a translating body.

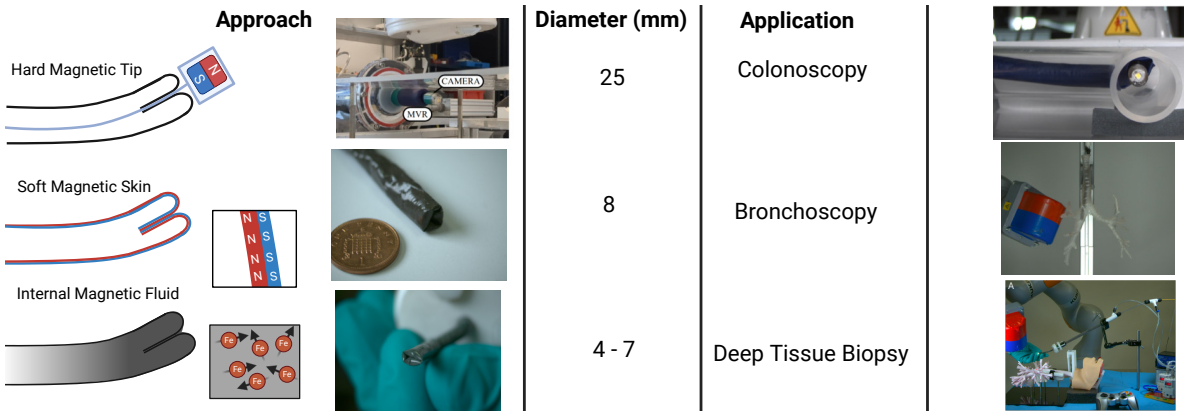


Fig. 2: Overview of the three discussed approaches for creation of magnetic vine robots.

B. Soft Magnetic Skin: Davy et al. [5]

The conventional magnetic tip design in vine robots often compromises a fully soft body, adding complexity to the miniaturization process. Our aim was to overcome this by developing a fully soft, magnetically-actuated vine robot, which is essential for navigating the intricate and delicate surgical environments ike the bronchial tree.

In collaboration with the Morimoto Lab at UCSD, we designed an 8 mm diameter robot characterised by a soft magnetic skin. This innovative design integrates magnetic micro-particles (NdFeB, 5 µm) within a thin silicone layer coated onto a pre-formed thin-wall tube. With careful consideration of the magnetisation direction, the robot preserves its fully soft architecture, embedding magnetic elements directly into the structure.

The robot's soft body and lack of rigid components enable it to navigate spaces narrower than its diameter. Experiments in navigating the upper bronchial tree using simple phantom models demonstrated its effectiveness in manoeuvring through complex pathways. While this fully soft and miniaturized design holds significant promise, it faces challenges with further miniaturisation. The necessity for a thin-wall structure inherently limits the magnetic volume, requiring strong external magnetic fields for effective control.

C. Internal Magnetic Fluid: Davy et al. [6]

Our most recent work has been motivated by these magnetic volume challenges inherent when aiming for more miniaturised designs whilst retaining the advantages of soft robots. To this end, we proposed a vine robot design based on internal magnetic fluid. This design, instead of internal pneumatic pressure, uses this active magnetic liquid as the source of pressurisation. Forces can then be induced on the fluid with magnetic gradient fields to steer the robot. This design, proves simpler to manufacture than the skin whilst significantly higher magnetic volume and therefore enabling simpler and more intutive operator control (see Figure 2).

We are now exploring magnetic fluid driven vine robot designs as small as 4 mm for various applications in the anatomy. One key application we envision is performing deep biopsies within the lungs.

CONCLUSION

Magnetic vine robots address the limitations of conventional endoscopic tools, such as low dexterity, limited control, and potential tissue trauma, by leveraging the unique properties of vine locomotion and magnetic actuation. By exploring three innovative approaches: hard magnetic tips, soft magnetic skins, and internal magnetic fluids, we have demonstrated the versatility of these systems for various endoluminal applications.

The distinct advantages of vine robots in terms of shear-free, tip focused movement enables them to navigate complex tortuous pathways. Via combination, with magnetic actuation, these robots are capable of precise control of the growth direction, enabling steering through bifurcations within the anatomy whilst minimising potentially traumatic tissue interaction. The three approaches outlined here each provide unique benefits and address specific surgical challenges. Future advancements will consider sensing methodologies within the robots for closed loop magnetic control, integrate diagnostic and therapeutic functions, and expand the use of these robots across a wider range of medical procedures.

REFERENCES

[1] S. Masuda, "Apparatus for feeding a flexible tube through a conduit," *GB-1534441-A, UK patent*, vol. 1, no. 6, 1978.

[2] P. Berthet-Rayne, S. H. Sadati, G. Petrou, N. Patel, S. Giannarou, D. R. Leff, and C. Bergeles, "Mammobot: A miniature steerable soft growing robot for early breast cancer detection," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5056–5063, 2021.

[3] N. G. Kim, N. J. Greenidge, J. Davy, S. Park, J. H. Chandler, J.-H. Ryu, and P. Valdastri, "External steering of vine robots via magnetic actuation," *Soft Robotics*, 2024.

[4] S. Al Harthy, H. Sadati, C. Girerd, S. Kim, Z. Wu, B. Saldarriaga, C. Seneci, T. Morimoto, and C. Bergeles, "Tip-growing robots: Design, theory, application," *IEEE TRANSACTIONS ON ROBOTICS*, 2024.

[5] J. Davy, N. Greenidge, S. Kim, L. J. Tinsley, P. Lloyd, J. H. Chandler, R. A. Harris, T. K. Morimoto, and P. Valdastri, "Vine robots with magnetic skin for surgical navigations," *IEEE Robotics and Automation Letters*, 2024.

[6] J. Davy, T. Dean, N. Greendige, B. Calmé, P. Lloyd, J. H. Chandler, and P. Valdastri, "Magnetic fluid driven vine robots for minimally invasive tissue biopsy sampling," *Advanced Intelligent Systems (Accepted)*, 2025.

Multi-fidelity model for thermal lesion prediction in robotic Ultrasound-Guided High-Intensity Focused Ultrasound surgery

F. Parrotta[1,2], G. Gamberini[1,2], S. Tognarelli[1,2], G. Stabile[1,2], A. Menciassi[1,2]

[1]*The BioRobotics Institute, Scuola Superiore Sant'Anna, 56127 Pisa, Italy.*
[2]*The Department of Excellence in Robotics and AI, Scuola Superiore Sant'Anna, 56127 Pisa, Italy.*
francesca.parrotta@santannapisa.it

INTRODUCTION

High-Intensity Focused Ultrasound (HIFU) thermal ablation relies on the localized deposition of acoustic energy to induce controlled tissue heating and coagulative necrosis, and its physical basis has been extensively characterized in both experimental and modeling studies [1], [2]. Numerical models based on acoustic propagation and bioheat transfer have been widely developed to reproduce tissue heating under HIFU exposure [3], [4], [5]. A well-established approach in HIFU thermal modeling is to compute tissue temperature through Pennes' equation, and to estimate thermal damage using dose-based criteria [6], [7]. In this framework, once the thermal field and model parameters are specified, the resulting lesion estimate is determined by the corresponding thermal dose distribution.

Although in vivo investigations remain essential for both model validation and clinical translation, the experimental characterization of HIFU lesion formation is inherently affected by variability, cost, and practical constraints. For this reason, in vitro and ex vivo approaches are commonly used as complementary models to reduce animal experimentation and enable more controlled and systematic studies [8]. At the same time, ex vivo tissues lack perfusion and convective heat loss, leading to lesion formation mechanisms that differ from those observed in vivo.

Multi-fidelity co-kriging is a probabilistic technique used in many different fields (e.g. geostatistics) that combines limited high-quality observations with more abundant lower fidelity measurements. This method has also been applied to physical systems governed by Partial Differential Equations (PDEs), including thermal fields [9], [10]. This rationale can be extended to HIFU lesion modeling, where high-fidelity experimental evidence is limited by cost and practical constraints, while lower-fidelity numerical or ex vivo data may be more readily available.

From this perspective, numerical simulations, ex vivo experiments, perfused tissue experiments, and vivo observations can be interpreted as correlated fidelity levels. The resulting lesion prediction would not be limited to a single deterministic contour but could instead be represented as a spatial probability field reflecting the experimental variability of HIFU-induced thermal damage. By integrating numerical predictions with progressively more realistic experimental evidence and explicitly quantifying uncertainty, this approach could support safer and more reliable HIFU treatment planning, particularly in autonomous surgical workflows.

This work was funded by "SURF-UP-Surgery by Using a Robotic Focused Ultrasound Platform" project under grant agreement no. FISA-2022-00770, within the Italina FISA programme—Fondo Italiano Scienze Applicate—HEALTH sector.

MATERIALS AND METHODS

The proposed framework was developed to model thermal lesion formation in robotic Ultrasound-Guided HIFU (USgHIFU) procedures by integrating data sources characterized by different levels of fidelity. A probabilistic multi-fidelity co-kriging approach based on Gaussian-process regression [11] was adopted to combine numerical simulations and experimental observations into a unified lesion prediction model.

Three fidelity levels were considered: (i) Low-fidelity (LF) data were obtained from numerical simulations of HIFU-induced thermal ablation performed in COMSOL Multiphysics [12]; (ii) Medium-fidelity (MF) data were obtained from ex vivo HIFU experiments performed on chicken breast tissue; (iii) High-fidelity (HF) data were acquired from artificially perfused ex vivo porcine kidney experiments designed to reproduce more realistic thermal transport conditions.

Artificial perfusion was introduced through a hydraulic circuit connected to the renal vasculature, with the inlet positioned at the renal artery and the outlet at the renal vein, as shown in Figure 1. Prior to HIFU exposure, the perfusion circuit was maintained active for 30 minutes to remove residual coagulated blood from the vascular network, while the pump was kept active during sonication to maintain active flow throughout the experiment. The experimental platform incorporates a 6-degree-of-freedom robotic manipulator (M0609, Doosan, Korea) holding the therapeutic HIFU transducer (Imasonic, Voray-sur-l'Ognon, France). driven by a high-power signal generator (Image Guided Therapy, Pessac, France), capable of delivering a maximum electrical power of 408 W. For real-time treatment monitoring, the HIFU transducer is integrated with a confocal micro-convex probe (MC10-4N12-A4, Telemed, Vilnius, Lithuania).

RESULTS

The co-kriging model produced probabilistic lesion predictions over the investigated treatment parameter space. Rather than providing a single deterministic lesion estimate, the model returned a predictive probability distribution at each query point, including both the expected lesion size and the confidence level. Preliminary results showed coherent spatial interpolation across different fidelity levels, with HF perfused-organ

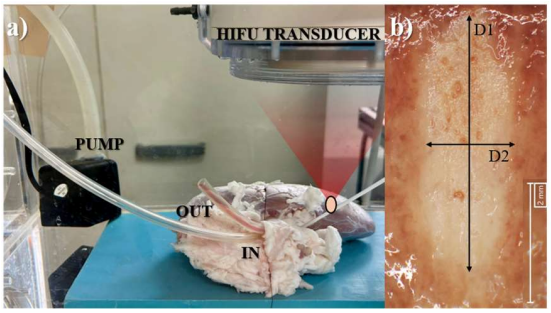


Figure 1. (a) Experimental robotic USgHIFU setup: the porcine kidney was connected to the artificial perfusion circuit through the renal artery (IN) and renal vein (OUT). (b) HRX-01 3D digital microscope (Hirox-USA) image of a thermal lesion generated in the perfused porcine kidney following HIFU exposure. D1 and D2 represent the major and minor axes of the ellipse-shaped HIFU thermal lesion.

data locally influencing lesion prediction and uncertainty estimates. In particular, regions supported by HF observations exhibited reduced prediction uncertainty, while sparsely sampled regions retained higher uncertainty levels.

This approach enabled simultaneous visualization of predicted lesion extent and confidence distribution, supporting uncertainty-aware interpretation of HIFU thermal damage. As shown in Figure 2 , D1 was markedly corrected across the different fidelity levels, whereas D2 exhibited more limited variations, suggesting either a lower sensitivity to perfusion effects or a more accurate description already provided by the numerical simulations.

CONCLUSION AND DISCUSSION

The proposed work introduces a probabilistic multi-fidelity framework for thermal lesion prediction in robotic USgHIFU procedures by combining numerical simulations, ex vivo experiments, and perfused-organ studies within a unified co-kriging model. The adopted strategy enables the integration of heterogeneous sources of information characterized by different levels of physiological realism, experimental complexity, and cost. By accounting for the uncertainty associated with lesion prediction and the experimental variability known to affect HIFU ablation studies, the framework supports safer and more standardized treatment planning within UsgHIFU workflows. Future work will focus on validating the proposed approach on larger experimental datasets and incorporating in vivo observations as the highest-fidelity source of information.

REFERENCES

[1] J. E. Kennedy, "High-intensity focused ultrasound in the treatment of solid tumours," *Nature Reviews Cancer, 2005.*

[2] G. ter Haar et al., "High intensity focused ultrasound: Physical principles and devices," *International Journal of Hyperthermia*, 2007.

[3] P. Montienthong et al., "Focused ultrasound ablation for the treatment of patients with localized deformed breast cancer: Computer simulation," *J. Heat Transfer*, 2020.

[4] K. P. Curra et al., "Numerical simulations of heating patterns and tissue temperature response due to high-intensity focused ultrasound," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 2000.

[5] A. Grisey et al., "Simulation of high-intensity focused ultrasound lesions in presence of boiling," *J. Ther. Ultrasound*, 2016.

[6] H. H. PENNES, "Analysis of Tissue and Arterial Blood Temperatures in the Resting Human Forearm," 1948.

[7] S. A. Sapareto et al., "Thermal dose determination in cancer therapy," *Int. J. Radiat. Oncol. Biol. Phys.*, 1984.

[8] A. Alassaf et al., "In vitro methods for evaluating therapeutic ultrasound exposures: Present-day models and future innovations," 2013.

[9] L. Parussini et al., "Multi-fidelity Gaussian process regression for prediction of random fields," *J. Comput. Phys.*, 2017.

[10] M. Guo et al., "Multi-fidelity regression using artificial neural networks: efficient approximation of parameter-dependent output quantities," 202.

[11] R. C. Kitson et al., "High-speed vehicle fluid-structure-jet interaction analysis and modeling," *58th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 2017.*

[12] F. Parrotta et al., "A Polynomial Model for Estimation of Ex-Vivo HIFU Thermal Lesion Dynamics Based on Pressure Amplitude and Sonication Time," *Applied Sciences, 2026.*

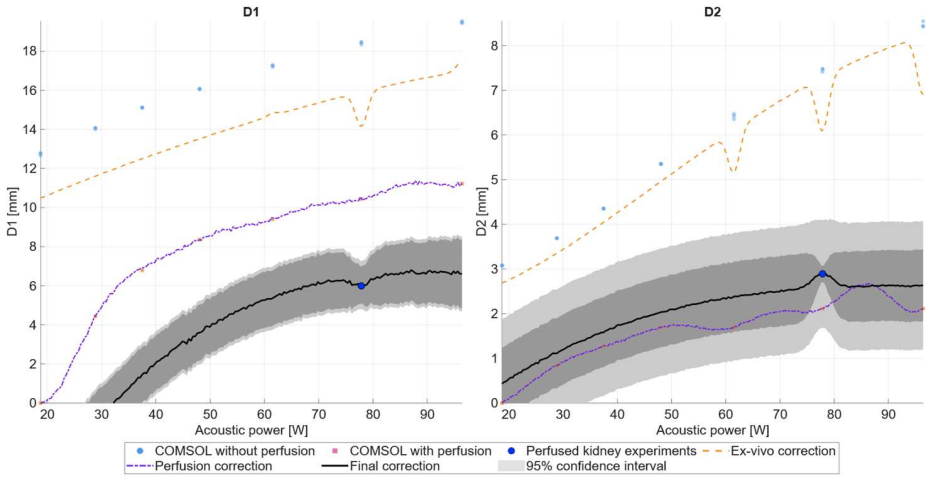


Figure 2: Multi-step multi-fidelity prediction of lesion dimensions D1 and D2. The COMSOL model without perfusion is sequentially corrected using ex-vivo data, perfusion-aware COMSOL simulations, and perfused organ experiments. The shaded grey region represents the 95% confidence interval of the final co-kriging output, allowing the model to reflect the experimental variability of HIFU lesion and to support uncertainty-aware interpretation of thermal damage.

Vision-Based Surgical Skill Assessment for Ring Transfer Tasks Using YOLO Segmentation

Sofia De Sanctis[1], Andrea Roberti[1], and Riccardo Muradore[1]

[1] *Department of for Innovation Medicine, University of Verona, Italy*

Introduction

Objective surgical skill assessment is becoming increasingly important in robot-assisted and minimally invasive surgery training [1]. Robot-assisted surgery requires advanced psychomotor abilities, including bimanual coordination, depth perception, camera control, and precise object manipulation. Conventional assessment methods are often subjective and rely on expert observation, limiting reproducibility and scalability. Peg transfer tasks are widely adopted in surgical training because they reproduce fundamental surgical gestures such as grasping, transferring, and precise placement [2], [3]. However, many existing assessment systems require either manual annotations, robotic kinematic data, or dedicated external tracking infrastructures, reducing their applicability in low-cost training environments. In this work, we propose a vision-based framework for automated surgical skill assessment using a multi-phase peg-and-ring transfer task inspired by Fundamentals of Laparoscopic Surgery (FLS) exercises [4]. The protocol consists of sequential phases in which colored rings are transferred between pegs while alternating the active surgical instrument. To increase task complexity, selected phases introduce controlled camera viewpoint changes, requiring the operator to adapt to modified visual conditions while maintaining accurate object manipulation.

The proposed framework employs a YOLO-based [5] instance segmentation model to detect and segment rings and pegs directly from monocular RGB video. The segmentation outputs are converted into object centroids and tracked over time to automatically reconstruct the task workflow. Based on the detected object interactions, the system extracts quantitative performance indicators including total task completion time, phase duration, transfer accuracy, motion continuity, ring drops, idle time, and bimanual coordination. The segmentation model was trained using manually annotated images acquired under different viewpoints and illumination conditions to improve robustness during camera motion.

Preliminary experiments demonstrate the feasibility of combining real-time object segmentation with automated workflow analysis using only monocular video recordings. Unlike approaches requiring robotic kinematics or external sensing systems, the proposed framework provides an inexpensive and scalable solution that relies exclusively on visual information, making it suitable for standard surgical training setups.

The main contributions of this work are:

- A multi-phase peg-and-ring transfer protocol incorporating camera viewpoint changes and instrument exchange;
- A YOLO-based instance segmentation pipeline for real-time detection and tracking of rings and pegs from monocular video;
- An automated workflow analysis framework for extracting quantitative surgical skill assessment metrics;
- A low-cost vision-only solution that does not require robotic kinematic information, wearable sensors, or external tracking systems.

Materials and Methods

Experimental Task

The proposed experimental setup is based on a peg-and-ring transfer task inspired by standard surgical training exercises. The task consists of multiple sequential phases in which the operator transfers colored rings among different pegs using two surgical instruments.

During the execution of the task, the active ring is exchanged between the two instruments at each phase transition. In addition, two phases include intentional camera viewpoint changes obtained by moving the camera around the workspace. These variations are introduced to evaluate the surgeon's adaptability to changing visual conditions and to increase the complexity of the manipulation task.

The setup includes:

- multiple colored rings,
- fixed pegs positioned inside the workspace,
- two robotic instruments (large needle driver),
- a monocular RGB camera for video acquisition.

All task executions are recorded as continuous video sequences and processed offline for automated analysis.

YOLO-Based Segmentation

To automatically detect the task elements, a YOLO-based instance segmentation framework was adopted. The model was trained to identify and segment both rings and pegs directly from monocular RGB frames (Figure 1). The training dataset consisted of 3000 manually annotated images acquired under different viewpoints and illumination conditions. Images were randomly divided into training, validation, and test sets (80%, 10%, and 10%, respectively). Standard data augmentation, including random scaling, rotation, and brightness variations, was applied to improve robustness against camera motion and illumination changes.

TABLE I

QUANTITATIVE COMPARISON BETWEEN THE EXPERT DVRK OPERATOR AND NOVICE USERS.

User Id	**Phase Times** (s)	**Total** (s)	**# Clutch**	**Time in Clutch** (s)	**Jerk** (m/s^3)	**Track Err.** $(pixel)$
User Expert	17 / 9 / 14 / 22 / 53	115	3	1.86	12.7	129
User 1	21 / 12 / 13 / 29 / 51	126	0	0	12.6	141
User 2	35 / 23 / 20 / 39 / 139	256	0	0	8.0	189
User 3	18 / 28 / 6 / 34 / 63.0	149	10	11.7	11.6	198
User 4	32 / 30 / 34 / 36 / 123	255	15	15.22	8.1	175
User 5	24 / 19 / 16 / 32 / 77	168	6	6.9	12.7	225
User 6	34 / 26 / 14 / 35 / 60	169	3	4.07	10.5	175

During inference, the segmentation pipeline provides pixel-wise masks for rings and pegs, from which object centroids are extracted and temporally associated to reconstruct the task workflow.



Fig. 1. Peg-and-ring experimental setup. Left: example frame from the monocular camera showing YOLO-based detection of pegs and rings. Right: corresponding view of the physical setup used for the task execution.

Skill Assessment Metrics

Based on the segmented video stream and dVRK setup, several quantitative metrics were extracted to objectively evaluate the operator's performance. The YOLO-based segmentation framework was also used to automatically identify the different task phases by analyzing the spatial and temporal relationships between rings, pegs, and surgical instruments during workflow execution. From this information, metrics such as total task completion time, number of successful transfers, failed placements and ring drops, instrument exchange accuracy, motion continuity, jerk, idle time, and performance variations after camera viewpoint changes were computed. These measurements provide an objective characterization of dexterity, bimanual coordination, and adaptation capabilities during the execution of the surgical task, while also enabling automatic workflow phase recognition directly from monocular video recordings. The proposed analysis pipeline processes each video frame using a YOLO-based detector to estimate the spatial position of rings and pegs throughout the task. Temporal consistency between consecutive detections was exploited to identify phase transitions, evaluate camera centering accuracy, and estimate motion stability during manipulation. The extracted trajectories further enabled the computation of motion smoothness indicators, clutch-related statistics, and camera tracking error metrics, which were automatically synchronized with the corresponding workflow phases for subsequent analysis.

RESULTS AND DISCUSSION

Preliminary experiments demonstrated consistent segmentation of rings and pegs throughout all task phases, including those involving camera motion and instrument exchange, enabling automatic workflow reconstruction from monocular video. A quantitative comparison was conducted between one expert dVRK operator and six novice users. Table I summarizes the extracted metrics, including phase execution times, total completion time, clutch usage, motion jerk, and camera tracking error.

The expert completed the task faster while exhibiting lower clutch usage and tracking error, indicating more efficient and stable performance. Novice users generally required longer execution times and showed greater variability across task phases, particularly following camera viewpoint changes. Differences in jerk and clutch statistics further reflected variations in motion smoothness and camera control strategies. These findings suggest that the proposed metrics are sensitive to differences in operator experience and can capture distinct aspects of surgical performance beyond task completion time alone. In particular, the combination of execution time, motion smoothness, clutch usage, and camera tracking error provides a more comprehensive characterization of technical proficiency. Although the present study represents a preliminary feasibility evaluation, the proposed framework shows potential for supporting objective and repeatable assessment in surgical training environments.

Overall, the proposed framework demonstrates the feasibility of objective surgical skill assessment using only monocular video. Future work will focus on larger-scale validation and the integration of surgical instrument tracking.

REFERENCES

[1] M. Levin, T. McKechnie, S. Khalid, T. P. Grantcharov, and M. Goldenberg, "Automated methods of technical skill assessment in surgery: a systematic review," *Journal of surgical education*, vol. 76, no. 6, pp. 1629–1639, 2019.

[2] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Autonomous sequential surgical skills assessment for the peg transfer task in a laparoscopic box-trainer system with three cameras," *Robotica*, vol. 41, no. 6, pp. 1837–1855, 2023.

[3] D. Meli, P. Fiorini, and M. Sridharan, "Towards inductive learning of surgical task knowledge: A preliminary case study of the peg transfer task," *Procedia Computer Science*, vol. 176, pp. 440–449, 2020.

[4] N. J. Soper and G. M. Fried, "The fundamentals of laparoscopic surgery: its time has come," *Bull Am Coll Surg*, vol. 93, no. 9, pp. 30–32, 2008.

[5] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of yolo algorithm developments," *Procedia computer science*, vol. 199, pp. 1066–1073, 2022.

Advanced 5G Standalone Infrastructure for Remote Surgical Robotics: Design and Validation of the OSLO Platform

J.A. Sánchez-Margallo[1], C. Argent Masa[2], F.M. Sánchez-Margallo[3]

[1]*Robotics, Image-Guided Surgery, and Bioprinting Unit, Jesús Usón Minimally Invasive Surgery Centre; jasanchez@ccmijesususon.com*

[2]*Bioengineering and Health Technologies Unit, Jesús Usón Minimally Invasive Surgery Centre; cargent@ccmijesususon.com*

[3]*Scientific Direction, Jesús Usón Minimally Invasive Surgery Centre; msanchez@ccmijesususon.com*

INTRODUCTION

Remote robotic surgery is considered one of the most demanding applications for next-generation communication systems due to the strict requirements for ultra-low latency, communication stability, high-bandwidth audiovisual transmission, and operational reliability [1]. Recent advances in 5G Standalone (SA) infrastructures and Open RAN architectures provide new opportunities for the deployment of flexible and low-latency communication frameworks suitable for surgical teleoperation [2]. However, the integration and validation of complete experimental ecosystems combining advanced 5G infrastructures, robotic control, real-time multimodal video transmission, and remote interaction remain limited.

The OSLO (remOte SurgicaL rObotics) project was developed to investigate the feasibility of advanced 5G communication infrastructures for teleoperated minimally invasive robotic surgery. The objective of this work was to design, deploy, optimize, and conduct a preliminary validation of a complete 5G-enabled robotic teleoperation platform integrating robotic manipulation, multimodal surgical video transmission, and remote audiovisual interaction.

MATERIALS AND METHODS

An intra-site experimental teleoperation platform was deployed at the Jesús Usón Minimally Invasive Surgery Centre (CCMIJU; Cáceres, Spain), with the robot and remote operator located in separate operating rooms. The final deployment comprised a private OpenAirInterface-based 5G SA network [3] with a shared 5G Core and Central Unit (CU) and two distributed unit/radio unit (DU/RU) branches equipped with USRP N310 radios. The two cells operated in bands n78 and n77 using 100-MHz TDD channels, 30-kHz subcarrier spacing, and 273 physical resource blocks (PRBs); the robot and operator computers were connected as separate user equipment (UEs) through Quectel modems. The system also integrated a KUKA LBR iiwa robotic arm controlled through a ROS-based application, integrated with a laparoscopic camera (Fig. 1), and the VEGA teleassistance platform (TedCas, Spain) for real-time multimodal video and audio transmission. The platform implemented unilateral position-command teleoperation with multimodal visual and audio feedback, but without force or haptic feedback.

The deployed network architecture incorporated dedicated communication pathways between the robotic system, the operating room (OR) environment, and the remote operator station. Validation activities were structured according to a dedicated verification and validation protocol covering both technical infrastructure verification and functional validation of the teleoperated surgical workflow.

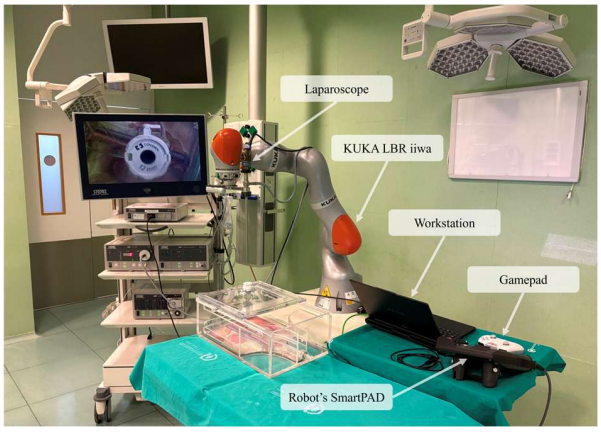


Figure. 1. Experimental setup of the surgical simulation environment, including the KUKA LBR iiwa collaborative manipulator equipped with a laparoscope that operates remotely on a laparoscopy simulator. Source: [4].

Laboratory validation studies were initially performed to verify the integration of the 5G Core Network, gNB infrastructure, radio-frequency communication components, and user equipment. Additional optimization studies evaluated configurable parameters associated with low-latency communications, including bandwidth adaptation, temporal slot allocation, multiplexing configurations, and URLLC-oriented service configurations.

Functional validation included assessment of: Remote robotic control; transmission of laparoscopic video; transmission of OR environmental video; transmission of

medical imaging (e.g. Ultrasound imaging); simultaneous integration of multiple audiovisual sources; P2P and Relay communication modes; stability and quality of remote communication.

RESULTS

The OSLO platform successfully demonstrated the feasibility of integrating a private 5G SA Open RAN infrastructure with a teleoperated robotic surgery environment. Stable bidirectional communication between the robotic system and the remote operator station was achieved using the deployed 5G infrastructure. In addition, the VEGA teleassistance environment enabled simultaneous real-time visualization and management of multiple surgical information sources, including laparoscopic imaging, OR cameras, robotic control interfaces, and additional medical imaging streams. Network measurements showed an average UE-to-UE round-trip time of approximately 17 ms, while the VEGA application reported communication latencies below 30 ms for P2P connectivity over the private 5G infrastructure. Relay connectivity through the VEGA cloud server resulted in an average latency of approximately 40 ms (Table 1).

Table 1. Latency and throughput performance of the OSLO teleoperation platform under different network configurations.

Measurement	Metric	Result
UE-to-UE over 5G	RTT	≈17 ms average
VEGA P2P over 5G	Application-reported latency	<30 ms
VEGA Relay	Host–server–remote latency	≈40 ms
OAIBOX throughput, simultaneous UEs	Download	251.7–263.8 Mbps
OAIBOX throughput, simultaneous UEs	Upload	62.9–70.3 Mbps

The validation studies additionally demonstrated the capability of the platform to support synchronized multimodal communication workflows required for remote surgical assistance and robotic teleoperation. Network optimization analyses identified several configurable parameters relevant for improving low-latency communications in surgical teleoperation applications, including bandwidth allocation, temporal slot configuration, and URLLC-oriented communication strategies.

CONCLUSION AND DISCUSSION

The results obtained in the OSLO project are consistent with the growing body of literature supporting the feasibility of 5G-enabled robotic telesurgery. Recent studies have highlighted that remote robotic surgery imposes stringent communication requirements in terms of latency, reliability, packet loss, and multimodal data transmission, particularly when real-time audiovisual and haptic feedback are involved [1]. In this context, the present work demonstrated the successful integration of a private 5G SA infrastructure with robotic teleoperation and multimodal surgical communication workflows.

Unlike previous telesurgery demonstrations using surgical-grade platforms and commercial networks, including a reported 18-ms command latency over 300 km [5], OSLO focuses on an intra-site, open, and fully configurable OpenAirInterface/Open RAN research environment.

This represents a relevant contribution toward reproducible and flexible research environments for future telesurgery investigations. Another relevant aspect of the present work is the evaluation of configurable network optimization strategies, including bandwidth adaptation, temporal slot allocation, multiplexing configurations, and URLLC-oriented communication parameters. The multimodal integration capabilities demonstrated in OSLO are also relevant in comparison with previous telesurgery platforms. Several recent studies have emphasized that future telesurgical systems must simultaneously support laparoscopic video, environmental imaging, robotic control signals, audio communication, and potentially haptic feedback streams under strict synchronization requirements [1].

The present study has several limitations. Validation was performed in an experimental environment, and no clinical procedures were conducted. In addition, quantitative characterization of packet loss, jitter, scalability under heavy network load, and long-distance communication performance requires further investigation. Overall, these results demonstrate the technical feasibility of integrating an open private 5G SA infrastructure with intra-site robotic camera teleoperation and multimodal communication. However, they do not establish clinical suitability or long-distance telesurgical performance.

ACKNOWLEDGEMENTS

This research was funded by the Spanish Ministry for Digital Transformation and Civil Service (TSI-064100-2022-025) and Junta de Extremadura (GR24210).

REFERENCES

[1] Dohler M, Saikali S, Gamal A, Moschovas MC, Patel V. "The crucial role of 5G, 6G, and fiber in robotic telesurgery". J Robot Surg; 2024;19(1):4.

[2] 3GPP. "Release 15–17 specifications for 5G communication systems." 3rd Generation Partnership Project (3GPP); 2018–2023.

[3] Kaltenberger F, Silva AP, Gosain A, Wang L, Nguyen TT. "OpenAirInterface: Democratizing innovation in the 5G Era." Computer Networks; 2020; 176:107284.

[4] Renedo-Alonso I, Sánchez-Margallo JA, Arana-Arexolaleiba N, Elguea-Aguinaco Í. "Adaptive Control of the Redundant Axis of a Surgical Robot for OR Workspace Optimization Using Reinforcement Learning". Sensors (Basel);2026; 5;26(9):2881.

[5] Xie X, Tian Y, Huang J, Luo Q, Chen T."Surgery without distance: will 5G-based robot-assisted telesurgery redefine modern surgery?" Transl Lung Cancer Res; 2025;14(5):1821-1829.

Evaluating localized muscle fatigue during suturing task on surgical training

A. Fernández-Montero[1], D. Caballero[1], M.J. Pérez-Salazar[1], J.A. Sánchez-Margallo[1], F.M. Sánchez-Margallo[2]

[1]Robotics, Bio-printing and Image-guided surgery Unit, [2]Scientific Direction

Jesús Usón Minimally Invasive Surgery Center, ES-10071, Cáceres, Spain.

{dcaballero, afernandez, jasanchez, msanchez}@ccmijesususon.com

INTRODUCTION

Minimally Invasive Surgery (MIS) has experienced significant growth in recent decades. The advantages of MIS for patients are well-documented and have been extensively discussed in scientific literature [1]. However, MIS procedures are technically intricate, resulting in an extensive and demanding learning curve. Consequently, comprehensive surgical training is essential for the different MIS specialties, techniques, and procedures that align with the training requirements and readily available resources.

In general, novice surgeons undergo a training process under the supervision of experienced surgeons, which enables them to evaluate their learning. The initial stages involve visualizing the various MIS procedures performed by experienced surgeons. Subsequently, novice surgeons are typically assigned to training on simulators, where they practice some surgical tasks as suturing.

The miniaturization of devices has enabled the integration of wearable technology into the surgical environment. Notably, systems for recording and analyzing muscular activity, particularly electromyography (EMG) systems, have been introduced [2].

Consequently, this study seeks to design and to evaluate muscular activity and localized muscle fatigue on the surgeon during suturing on surgical training in conventional laparoscopy and Robot-assisted surgery (RAS), implementing a set of wearable devices, preventing ergonomic deficiencies and improving the surgical quality.

MATERIALS AND METHODS

Experimental setup

The data for this study were collected during forty-two surgical training sessions, including different surgical training tasks such as eye-hand coordination, cutting and suturing tasks. Nevertheless, only suturing was evaluated in the current study. Half of these training sessions (n = 21) were performed by using conventional laparoscopy techniques and the remaining techniques using RAS. The surgeons were two women and five men between the ages of 25 and 45, all of whom were right-handed. The data were recorded using a portable, wireless EMG system that allowed the EMG signal to be recorded for subsequent calculation of localized muscle fatigue in eight muscle groups bilaterally.

Laparoscopy surgery sessions

Laparoscopy surgery sessions were carried out with the Olympus® Visera III Tower (Olympus®, Tokyo, Japan) and Karl Storz® surgical instruments were used during laparoscopy sessions. All surgical tasks were carried out in a laparoscopic training simulator.

Robot-assisted surgery sessions

RAS sessions were carried out with the Versius® robotic platform (CMR Surgical®, Cambridge, United Kingdom) was used. As with laparoscopic technique, all surgical tasks were carried out in a surgical training simulator.

Suturing tasks

Surgeons were asked to perform a suture on a previous dissection in simulator. Suturing consisted of two simple sutures and one double suture in opposite directions [3]. A needle holder was used in each hand, with a time limit of ten minutes.

Wearable device

The wireless EMG TRIGNO® Avanti system from DELSYS® was used to record the muscle activity by means of EMG signals. This system is composed by sixteen sensors with a sampling rate of 1024 Hz. The EMG signal of following muscle groups was recorded bilaterally: upper trapezius (UP_TRAP), middle trapezius (MID_TRAP), erector spinae (ER_SPIN), deltoids (DEL), brachioradialis (BRACH), triceps brachii (TRI_BRA), vastus lateralis (VAS_LAT), and gastrocnemius medialis (GAS_MED). The EMG sensors were placed on each muscle according to the SENIAM guidelines [4].

Localized muscle fatigue

The EMG values were presented as percentage of maximum voluntary contraction (%MVC). This was obtained separately for each muscle group just before each test by asking each surgeon to perform specific contractions against a fixed resistance in order to fix the

performance of the baseline. For the study of localized muscle fatigue, the evolution of the spectral amplitude and median frequency of the muscle activity was analyzed. Muscle fatigue was quantified using Joint Analysis Spectral and Amplitude (JASA) method [5].

RESULTS

Evaluation of localized muscle fatigue

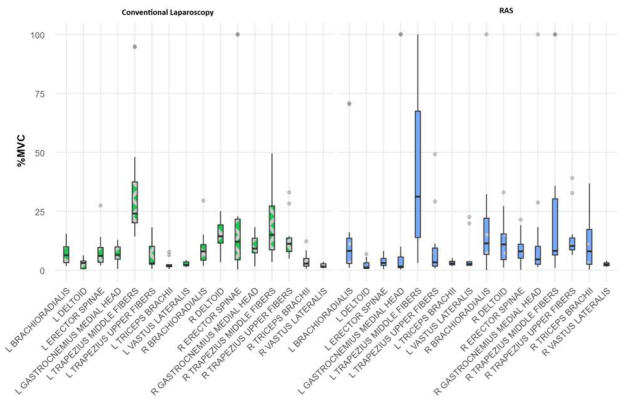


Figure. 1. Comparison of muscle activity (%MVC) of surgeons during performance of conventional laparoscopy (green) and RAS (blue).

In general, the surgeons had less muscle activity in the left-sided muscle than the right-sided except for the trapezius middle (Figure 1). This fact could be related with the dominant side of all surgeons (right side). Consistently, muscle activity on the dominant side also appears to increase slightly over time compared with the non-dominant side, suggesting that the dominant side may be more susceptible to developing musculoskeletal disorders [6]. This fact demonstrates the different use of muscle between conventional laparoscopy and RAS. All muscles bilaterally showed statistically significant differences between conventional laparoscopy and RAS ($p<0.001$) and between dominant and non-dominant side ($p<0.001$).

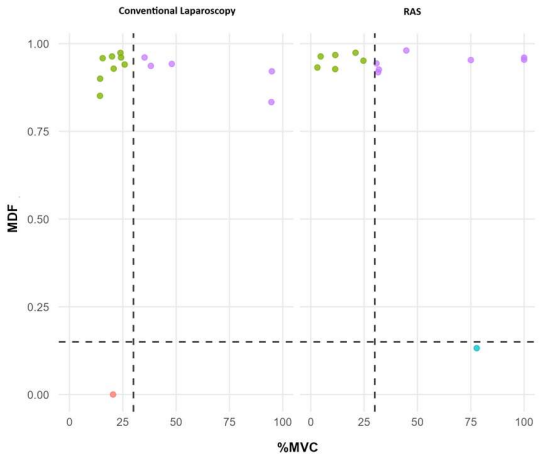


Figure. 2. Comparison of fatigue and muscle strength increase/decrease of left trapezius middle for surgeons during conventional laparoscopy (left) and RAS (right). Low %MVC and High MDF (Muscle recovery – Red color). Low %MVC and Low MDF (Force decrease – Green color). High %MVC and Low MDF (Muscle fatigue – Violet color). High %MVC and High MDF (Force increase – Blue color).

Analyzing the localized muscle fatigue during the surgical training, the different surgeons showed lower muscle fatigue and higher force use in RAS than in conventional laparoscopy (Figure 2).
Muscle activity is closely related to muscle fatigue and prolonged muscle activity of small, single muscle fibers can cause degenerative muscles changes even with very low levels of muscle activity. For this reason, ergonomic positioning seems highly crucial during MIS activities. The muscles analyzed in this study have been explored throughout other studies on MIS, with the main muscles identified being: the upper and the middle trapezius and the brachioradialis [6]. As example, localized muscle fatigue for left middle trapezius is showed on Figure 2.

CONCLUSIONS

This research contributes to the understanding of ergonomic risks in surgical training, presenting a substantial advance in the integration of wearable devices and the evaluation of the localized muscle fatigue. These results have the potential to significantly improve surgical training programs, the design of ergonomic surgical instruments and the overall well-being of surgeons.

Acknowledgments

This work was co-funded by the European Union through the Interreg VI-A Spain–Portugal (POCTEP) 2021–2027 Programme and the Junta de Extremadura (GR24210).

REFERENCES

[1] Gabrielson AT, Clifton MM, *et al.*, "Surgical ergonomics for urologists: A practical guide." Nat Rev Urol; 2021; 18:160-69.

[2] Merbah J, Caré BR, *et al.*, "A new approach to quantifying muscular fatigue using wearable EMG sensors during surgery: An ergonomic case study." Sensors (Basel); 2023; 23:e1686.

[3] Phan PT, Hoang TT, et al., "Smart surgical sutures using soft artificial muscles." Sci. Rep.; 2021; 11(1):22420.

[4] Hermens HJ, Freriks B, *et al.*, "Development of recommendations for SEMG sensor and sensor placement procedures" J Electr Kines; 2000; 10:361-74.

[5] Dufaug A, Barthod C, *et al.*, "New joint analysis of electromyography spectrum and amplitude-based methods towards real-time muscular fatigue evaluation during a simulated surgical procedure: A pilot analysis on the statistical significance" Med Eng & Phys; 2020; 79:1-9.

[6] Caballero D, Pérez-Salazar MJ, *et al.*, "Localized Muscular Fatigue in Robotic-Assisted Laparoscopic surgery: Predictive modeling study" JMIR Form Res; 2025; 9:e68536.

Preliminary Characterisation of Angular Overshoot and Oscillation in a Pneumatic Soft Continuum Robot

Q. Wei[1], H. Sun[1], M. Chauhan[1,*]
[1]*School of Physics, Engineering and Technology, University of York, York, United Kingdom*
*wsp525@york.ac.uk; ddf538@york.ac.uk; *manish.chauhan@york.ac.uk*

INTRODUCTION

Soft continuum robots are promising for minimally invasive navigation because their compliant bodies can deform continuously and distribute interaction forces over larger contact areas than rigid instruments [1, 2]. Pneumatic actuation is attractive because pressurised fluid can produce large deformation with simple, compliant structures [3].

However, pneumatic soft robots remain difficult to control accurately. Their response is affected by nonlinear material behaviour, geometry, pressure-line delay, air compressibility, and viscoelasticity [2,4]. Existing strategies for reducing transient error include feedback control and feedforward input shaping [5,6], but the transient features requiring suppression must first be identified.

Angular overshoot—the difference between maximum bending angle and final steady-state angle—is a common transient metric. However, a response with small overshoot may still exhibit post-rise oscillation, increasing tip-motion uncertainty and swept motion near anatomical boundaries. This work presents a preliminary experimental characterisation of single-channel transient bending, highlighting two distinct behaviours: overshoot-dominated and oscillation-dominated response.

MATERIALS AND METHODS

The experimental platform was a four-channel pneumatic soft continuum robot fabricated from silicone elastomer. Only chamber 1 was pressurised to isolate single-channel bending. Pressure commands were generated using a proportional pneumatic control system (SMC solenoid valve manifold with TENMA DC power supply), and pressure-ramp inputs were applied. Robot motion was recorded using an NDI Aurora electromagnetic tracking system, and bending angle was extracted from the measured tip trajectory. The experimental setup is shown in Fig. 1.

The bending angle response $\theta(t)$ was analyzed for steady-state angle θ_{ss} (from the final stable portion) and maximum angle θ_{max}. Angular overshoot was defined as $OS = \theta_{max} - \theta_{ss}$, and relative overshoot as $OS_{\%} = (\theta_{max} - \theta_{ss})/\theta_{ss} \times 100\%$.

Post-rise oscillation was characterized using FFT analysis. Trials were classified as overshoot-dominated when $OS_{\%} \geq 5\%$ and oscillation-dominated when $OS_{\%} < 5\%$ with post-rise oscillation amplitude $\geq 1^{\circ}$. These thresholds were chosen empirically to separate distinct behaviours in preliminary trials.

RESULTS

Figure 2 compares two representative single-channel transient responses. The pressure amplitude and ramp duration were intentionally varied together to explore the operational parameter space efficiently; this is therefore not a controlled one-factor study. The two cases were selected to contrast overshoot-dominated versus oscillation-dominated behaviours, demonstrating that peak overshoot and post-rise oscillation are related but not equivalent dynamic features.

In the low-to-moderate pressure case (Fig. 2a), steady-state was $\theta_{ss} \approx 8.73^{\circ}$ with maximum $\theta_{max} \approx 9.38^{\circ}$, giving overshoot of 0.66° (7.54%)—classified as overshoot-dominated. FFT shows dominant frequency 7.60 Hz.

In the higher-pressure case (Fig. 2b), steady-state was $\theta_{ss} \approx 41.80^{\circ}$ with maximum $\theta_{max} \approx 42.42^{\circ}$, giving overshoot of 0.62° (1.48%). With post-rise oscillation amplitude $\approx 1.50^{\circ}$, this met the criterion for oscillation-dominated behaviour. FFT shows dominant frequency 8.93 Hz.

The comparison suggests that angular overshoot alone may underestimate dynamic motion uncertainty. In confined navigation, residual oscillation can increase the swept envelope of the distal segment even when the peak angular overshoot is small. Therefore, transient safety evaluation should consider both target-angle exceedance and post-rise oscillation.

CONCLUSION AND DISCUSSION

This preliminary study demonstrates that peak angular overshoot is an incomplete metric for assessing transient safety of pneumatic soft continuum robots. While the higher-pressure case exhibited minimal relative overshoot (1.48%), it maintained significant post-rise oscillation ($\approx 1.50^{\circ}$), which increases swept volume during interventional procedures. For surgeons performing minimally invasive navigation in confined anatomical spaces, this distinction is critical: a robot with low overshoot but sustained oscillation may still increase the risk of unintended contact with surrounding tissue, which overshoot metrics alone would not capture.

The characterization reveals two distinct dynamic regimes: overshoot-dominated and oscillation-dominated responses. Small values of $\theta_{max} - \theta_{ss}$ do not guarantee smooth motion, as dominant frequency components (7.60–8.93 Hz) persist even when overshoot is suppressed. This underscores the necessity of multi-objective transient metrics accounting for both exceedance and oscillation when evaluating control strate-

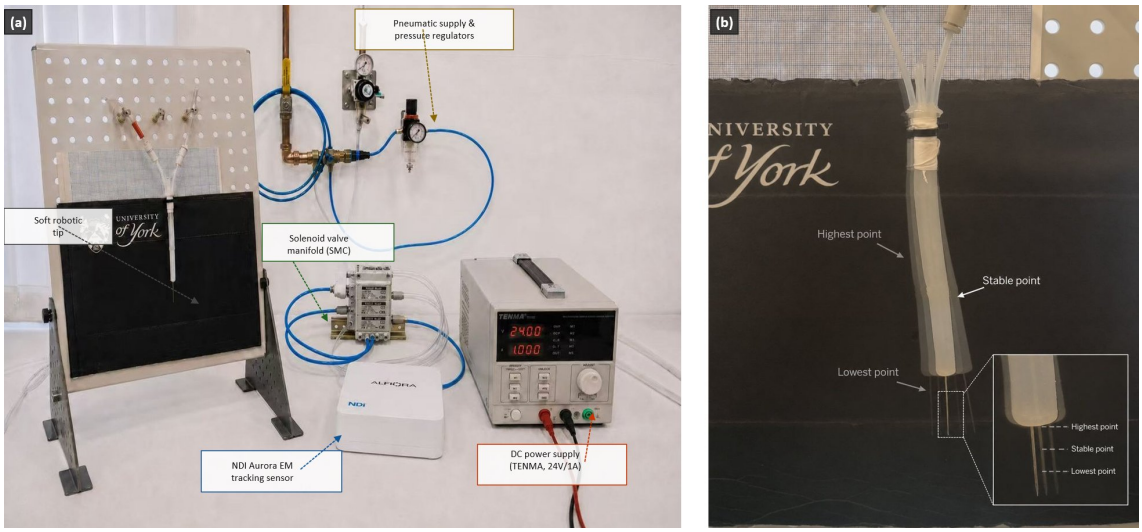


Figure 1. Experimental setup. (a) Benchtop overview showing pneumatic supply with pressure regulators, SMC solenoid valve manifold, TENMA DC power supply, NDI Aurora electromagnetic tracking system, and soft robotic tip mounted vertically on an adjustable stand. (b) Close-up view of the soft robotic tip.

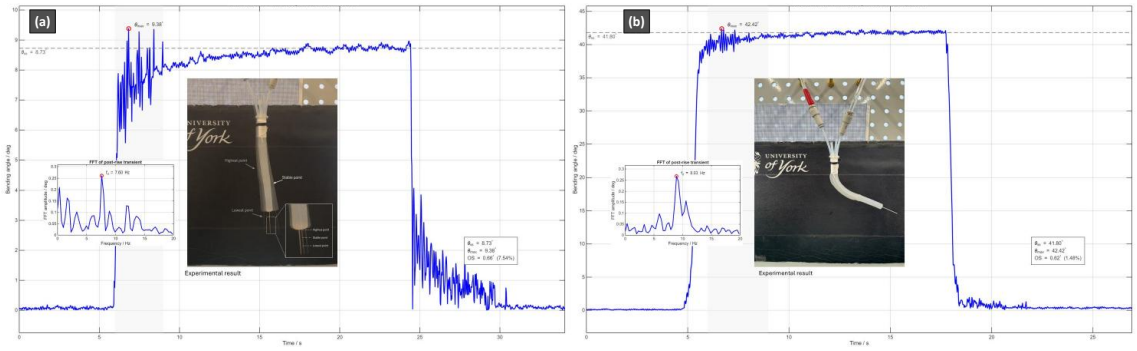


Figure 2. Representative transient responses: (a)Overshoot-dominated response, 0.7 bar /200 ms, 0.7 bar/200 ms; (b) Oscillation-dominated response, 1.0 bar /1000 ms, 1.0 bar/1000 ms.

gies for surgical robotics.

Future work will evaluate input-shaping and closed-loop compensation to suppress oscillations. The dataset will expand to generate a full pressure-ramp response map, facilitating feedforward command shaping for real-time surgical tasks. Dual-channel compensation will enable active damping, advancing toward closed-loop control ensuring both accuracy and safety in confined surgical environments.

REFERENCES

[1] C. Majidi, "Soft robotics: A perspective—current trends and prospects for the future," *Soft Robotics*, vol. 1, no. 1, pp. 5–11, 2014.

[2] T. Ashuri, A. Armani, R. J. Hamidi, T. Reasnor, S. Ahmadi, and K. Iqbal, "Biomedical soft robots: Current status and perspective," *Biomedical Engineering Letters*, vol. 10, pp. 369–385, 2020.

[3] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," *Nature*, vol. 521, pp. 467–475, 2015.

[4] R. J. Webster III and B. A. Jones, "Design and kinematic modeling of constant curvature continuum robots: A review," *International Journal of Robotics Research*, vol. 29, no. 13, pp. 1661–1683, 2010.

[5] N. C. Singer and W. P. Seering, "Preshaping command inputs to reduce system vibration," *Journal of Dynamic Systems, Measurement, and Control*, vol. 112, no. 1, pp. 76–82, 1990.

[6] W. E. Singhose, "Command shaping for flexible systems: A review of the first 50 years," *International Journal of Precision Engineering and Manufacturing*, vol. 10, no. 4, pp. 153–168, 2009.

A Low-Power Multi-Sensor Navigation Architecture for Spatial Mapping of Endoscopic Capsules

Sahil Anchal, Sean Coombes, Daragh Crowley,
Eoin Higgins, Herman Alexander Jaeger, and Pádraig Cantillon-Murphy

School of Engineering and Tyndall National Institute, University College Cork, Cork, Ireland

I. INTRODUCTION

Capsule Endoscopy (CE) is a minimally invasive diagnostic procedure in which a patient swallows a pill-sized device equipped with a microcamera, a battery, and a wireless transmitter[1], [2]. As the capsule is propelled naturally by gastrointestinal(GI) peristalsis, it captures high-resolution imagery of the deep small bowel and colon—areas traditionally difficult to access via conventional methods. Consequently, CE has gained significant traction among clinicians and patients alike. It offers a diagnostic alternative to traditional colonoscopy that eliminates the need for sedation, minimises physical discomfort, and provides a comprehensive view of the GI tract without the risks associated with invasive insertion.

However, the clinical utility of CE is significantly constrained by the absence of precise navigational capabilities. Despite the high diagnostic yield of the captured imagery, the 'blind' transit of the device means that detected pathologies often lack spatial context. Without the ability to correlate images with specific anatomical landmarks, the exact coordinates of a lesion within the convoluted GI tract remain unknown. This spatial ambiguity makes follow-up surgical or therapeutic interventions difficult, as clinicians cannot reliably return to the site of interest.

Electromagnetic tracking (EMT) is widely recognised as the clinical gold standard for minimally invasive surgical navigation, offering highly accurate, real-time spatial estimates in environments where optical line-of-sight is obstructed [3]. However, translating EMT to capsule endoscopy (CE) remains challenging due to the stringent power constraints of wireless nodes [1].

In this work, we designed an energy-efficient, EMT-enabled hardware platform designed for high-precision CE localisation. To address the inherent limitations of EMT, we incorporate a low-power Inertial Measurement Unit (IMU) to facilitate sensor fusion [4]. This approach enables a deliberate trade-off between localisation update frequency and battery longevity, which is critical for long-duration CE procedures [1]. We provide a systematic characterisation of the node's energy footprint across various operational states, establishing a baseline architecture for future power-optimised, long-term endoscopic tracking systems.

MATERIALS AND METHODS

The proposed system is structured as an integrated, multi-sensor wireless node designed to address the challenges of clinical localisation in capsule endoscopy (CE). As illustrated in Fig. 1a, the system architecture is based on the AnserEMT open-source research platform [3]. It comprises a field generator (FG) that produces a time-varying magnetic field, an EMT-navigated wireless capsule, and a visualisation interface for clinical monitoring. The FG is embedded in a wearable belt [2] worn by the patient, which receives in vivo endoscopic imagery and transmits it over a wireless network for real-time visualisation.

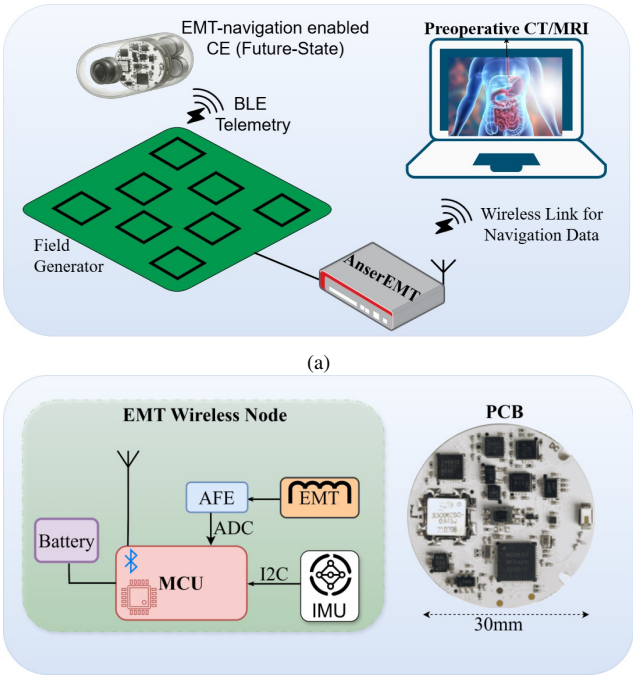


Fig. 1: EMT-Navigation enabled multi-sensor system for CE: (a) System level overview showing the capsule, field generator, and wireless data flow; (b) Functional block diagram and PCB prototype of the EMT wireless node [5]

The core of our designed system is in the EMT-enabled wireless node, illustrated in Fig. 1b. This node is designed around a power-efficient microcontroller unit (MCU:nrf52832) that interfaces with both the EMT sensor coil via an Analogue Front End (AFE) and an Analogue-to-Digital Converter (ADC), and with a low-power IMU via an Inter-Integrated Circuit (I2C) peripheral. The wireless sensor node is designed for edge-based signal pre-

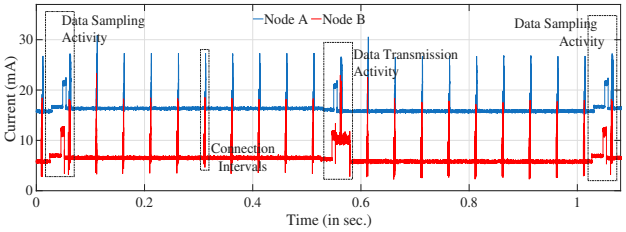


Fig. 2: Current consumption profile of wireless nodes.

processing rather than full-scale localisation. The capsule node utilises an on-board Discrete Fourier Transform (DFT) algorithm to directly extract magnitude and phase information from the EMT sensor coil signal [5]. This preprocessed data, alongside raw accelerometer data from the integrated IMU, is then transmitted via a low-power Bluetooth Low Energy (BLE) link to the wearable sensor belt. The sensor fusion algorithm, which calculates the final position and orientation estimates, resides on the wearable [4]. We utilised a real-time operating system (RTOS) based framework to manage the strict timing requirements of multi-modal data acquisition from IMU and EMT sensors. The firmware efficiently schedules preprocessing tasks and data transmission while maintaining the flexibility to adjust Output Data Rates. This capability allows the node to tune its performance to match the specific needs of the sensor-fusion algorithms on the belt, balancing tracking accuracy with energy efficiency. The current design iteration builds on our previous works [5], [6], featuring a compact 30 mm diameter form factor. While this size is suitable for bench-top validation and characterisation, our ongoing research is focused on shrinking the component footprint and optimising board-level integration, ensuring the system can be fully encapsulated within a standard commercial-sized capsule endoscope for clinical deployment.

DISCUSSION AND CONCLUSIONS

To validate the energy efficiency of the proposed architecture, we performed a systematic power characterisation of two hardware iterations using the Nordic Power Profiler Kit (PPK II) and provided a 3.3 volts from it. Node A, which utilised an external ADC interfaced alongside an AFE, exhibited an average current consumption of ∼14.5 mA due to significant peripheral and bus overhead, refer to Fig. 2. In contrast, using the onboard ADC instead of external components allowed Node B to consume ∼55% less energy, keeping its current consumption stable at ∼6.45 mA (refer to Fig. 2). To validate 3D spatial tracking, a zig-zag trajectory spanning a $30\times30\times30\,\mathrm{cm}^3$ active volume [5] was benchmarked using the wired interface, refer Fig. 3. The direct EMT position estimates demonstrated strong path fidelity, achieving a mean spatial trueness of ∼ 3.0 mm relative to the robotic ground truth.

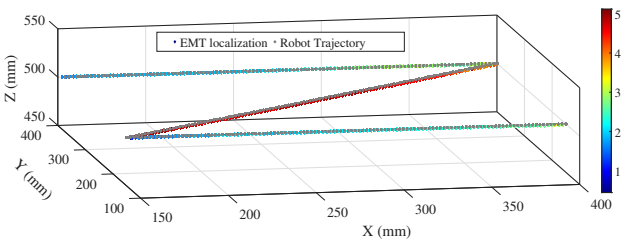


Fig. 3: Visualization of the EMT localization vs Robot trajectory

Standard commercial endoscopic capsules typically utilize batteries with capacities ranging from 55–60 mAh [2]. Based on these industry standards, the optimised Node B architecture can sustain approximately 9.3 hours of continuous operation, effectively meeting the 8-hour clinical requirement for a complete gastrointestinal transit study. Location-enabled capsules significantly enhance diagnostic precision by providing precise spatial coordinates for clinical findings, moving beyond unreliable time-based estimates. Integrated with IoT frameworks, the presented architecture enables remote patient monitoring, enabling hospital-grade diagnostics in a home setting. Furthermore, providing precise navigation data as input to Computer Vision algorithms enables automated reconstruction of a 3D gut profile. This spatial awareness improves image stitching and lesion mapping, creating a more effective, data-driven diagnostic tool for personalised medicine.

Acknowledgment: This work is supported by the European Union's Horizon Europe Excellent Science program under the Marie Sklodowska-Curie Actions Grant Agreement(Grant Agreement No 101081457) and in part from CÚRAM, Research Ireland Centre for Medical Devices [under Grant Reference 13/RC/2073_P2] and the ERC Proof of Concept Award 101189344 (2024).

REFERENCES

[1] M. A. Ali, N. Tom, F. N. Alsunaydih, and M. R. Yuce, "Recent advancements in localization technologies for wireless capsule endoscopy: A technical review," *Sensors*, vol. 25, no. 1, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/25/1/253

[2] I. M. Mehedi, K. P. Rao, F. M. Alotaibi, and H. M. Alkanfery, "Intelligent wireless capsule endoscopy for the diagnosis of gastrointestinal diseases," *Diagnostics*, vol. 13, no. 8, 2023. [Online]. Available: https://www.mdpi.com/2075-4418/13/8/1445

[3] H. Alexander Jaeger, A. M. Franz, O. Kilian, S. Alexander, T. Fabian, M.-H. Lena, and P. Cantillon-Murphy, "AnserEMT: the first open-source electromagnetic tracking platform for image-guided interventions," *Int. J. of Comp. Assisted Radiology and Surgery*, vol. 12, pp. 1056–1067, 2017.

[4] K. O'Donoghue, H. A. Jaeger, and P. C. Murphy, "Sensor fusion hardware platform for robust electromagnetic navigation," in *2022 IEEE Int. Inst. Meas. Tech. Conf. (I2MTC)*, 2022, pp. 1–6.

[5] D. Crowley, M. Cavaliere, E. Higgins, C. Pande, H. Alexander Jaeger, K. O'Donoghue, and P. Cantillon-Murphy, "Wireless electromagnetic sensors for image-guided cardiothoracic procedures," *IEEE Sens. J.*, vol. 24, no. 22, pp. 37 724–37 733, 2024.

[6] S. Coombes, E. Higgins, H. Jaeger, and P. Cantillon-Murphy, "Wireless electromagnetic tracking in capsule endoscopy using tmr sensing," in *2022 IEEE Int. Inst. Meas. Tech. Conf. (I2MTC)*, 2026, pp. 1–6.

Constraint Programming for Robotic Arm–Assisted Delivery of Ultrasound-Guided High-Intensity Focused Ultrasound (USgHIFU)

Manisha Waterston[1,2], Lewis Howell[1,2], Bailey Andrew[5], James Chandler[2], Tze Min Wah[3,4], and James McLaughlan[2,4]

[1]*School of Computing, University of Leeds*
[2]*School of Electronic and Electrical Engineering, University of Leeds*
[3]*Department of Radiology, Leeds Teaching Hospital Trust*
[4]*Faculty of Medicine and Health, University of Leeds*
[5]*School of Medical Sciences, University of Manchester*

INTRODUCTION

High Intensity Focused Ultrasound (HIFU) is a non-invasive and non-ionising treatment approach that uses high-amplitude ultrasound waves to thermally ablate the treatment area [1]. HIFU has been applied to a range of medical conditions, including cancers of the liver, breast, pancreas, and brain, as well as in benign tumours of the thyroid and prostate [2]. As HIFU delivers local heating (>55°C for longer than a second to induce cell death) in small volume doses (thermal lesions), multiple spatially distributed iterations are typically required to cover the full size of the treatment area [3].

To aid in lesion placement, HIFU can be delivered with Ultrasound guidance (USgHIFU), which combines real-time diagnostic imaging with treatment [4]. Integrating USgHIFU with co-robotic assistance can further improve treatment quality by offering precise automated beam positioning to minimise damage to surrounding healthy tissue [5]. However, to prevent healthy tissue from overheating between the point of entry and the target area, cooling intervals between each HIFU dosage are required [6]. This can extend treatment delivery, requiring patients to remain stationary for several hours, limiting the uptake of HIFU [2]. Current treatment planning methods rely on a clinician to highlight a planning target volume, with subsequent treatment lesions typically delivered in a raster pattern over the treatment area [7]; without consideration of acoustic propagation of the beam or procedural efficiency. Therefore, new planning techniques are needed to shorten treatment and provide more precise dosages that protect healthy tissue.

This paper proposes and evaluates a constraint programming approach to optimise HIFU lesion placement and ordering in co-robot assisted USgHIFU. The proposed approach is considered for HIFU treatment of thyroid nodules using a co-robotic system and based on ultrasound images, as shown in Figure 1. This aims to maximise the benefits of co-robot integration in USgHIFU through enhanced efficiency and precision of HIFU delivery, with the ultimate goal of facilitating broader clinical implementation.

MATERIALS AND METHODS

Two-dimensional (2D) B-mode ultrasound scans of thyroid nodules were analysed and extracted from a Leeds Teaching Hospitals dataset. A custom pixel measurement tool was developed to quantify nodule dimensions and define the acoustic pathway and tissue layer thicknesses directly from the scans. These parameters were used in an acoustic and thermal model (MATLAB k-Wave toolbox, Mathworks, USA) to simulate HIFU-based thermal ablation based on a 2 MHz HIFU wave with 0.5 MPa pressure [8]. Three thyroid nodules were selected from the dataset and simulated. The resulting thermal lesion sizes were extracted at the bottom, middle and top of each nodule to account for propagation path within the planning framework.

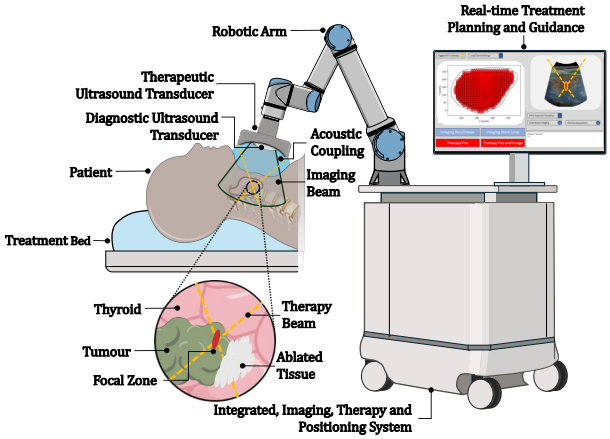


Fig. 1: Illustration of the USgHIFU setup with co-robotic delivery and HIFU delivery planning based on ultrasound imaging.

These lesion dimensions were incorporated into a constraint-programming-based treatment planning framework, using either region-specific lesion sizes or an average lesion size with region-specific heating and cooling times. Candidate lesion centres were generated using a 2D grid within the nodule boundary and encoded in a binary coverage matrix, where rows represent candidate lesions and columns represent grid points. A Google OR-Tools CP-SAT model formulated lesion selection as a binary optimisation problem and selected the minimum number of lesions required to achieve complete target coverage. An $(N + 2) \times (N + 2)$ binary overlap matrix was then constructed to encode geometric interactions between selected lesions, with additional rows and columns representing

fixed start and end nodes. Entries were assigned when lesions overlapped or touched, producing a symmetric adjacency matrix. This matrix was used in a second CP-SAT routing and scheduling model to optimise lesion ordering while satisfying geometric constraints, including a bottom-to-top progression, and reducing the risk of overtreatment in previously treated areas. Three 2D treatment plans were subsequently developed for implementation. An example treatment plan is shown in Figure 2 *a*). In this case, an average lesion size was assumed across the nodule, resulting in 114 lesions, each measuring 1.32×3.68 mm, to achieve complete ablation of the target region. A procedure time of 12 minutes was predicated, which is significantly faster than a raster-based approach which required 151 minutes as seen in Figure 2 *b*). The raster plan resulted in 98.75% tumour coverage, 71.36% overlap, and 14.55% healthy tissue ablation, with one lesion extending beyond the planning tumour boundary. In contrast, the CP framework achieved 99.85% tumour coverage while reducing overlap to 39.99% and healthy tissue ablation to 13.45%.

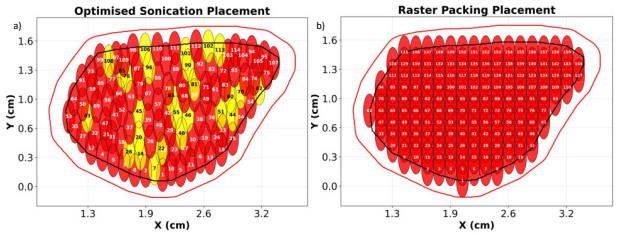


Fig. 2: a) illustrates the thyroid nodule, with the nodule boundary (black contour) and planning boundary (red contour). Red lesions indicate locations that can be treated consecutively without interruption, while yellow lesions require an enforced cooling interval before subsequent lesions. (b) The same thyroid nodule presented in (a), planned using a raster-based approach to lesion placement.

To deliver the treatment plans, an H102 Sonic Concepts transducer and ultrasound imaging probe will be integrated with a six-degree-of-freedom collaborative robot (UR3e, Universal Robots). The robotic platform utilises ROS2, Universal Robots drivers, PLUS, 3D Slicer, and MoveIt for pose estimation, visualisation, motion planning, and collision avoidance [9], [10], [11], [12]. A Python script has been developed to read the plans and coordinate robotic motion sequentially through the planned treatment points and trigger the HIFU transducer for the specified durations.

CONCLUSIONS AND DISCUSSION

This work presents an optimised, automated treatment planning framework for thyroid nodules based on 2D B-mode ultrasound images. While absolute optimality cannot always be guaranteed, the proposed method optimises lesion placement and ordering, while enabling real-time monitoring, guidance, and the delivery of co-robotic therapy. It is able to accommodate arbitrary nodule geometries, lesion sizes and different transducer characteristics, offering a practical balance between optimisation quality and applicability. The framework is limited by reduced optimisation quality for large nodules, simplified lesion modelling due to simulation constraints and idealised shapes and reliance on 2D analysis, though it remains adaptable and clinically practical. By utilising nodule specific geometries based on ultrasound images, optimising treatment plans and co-robotic assisted delivery, the proposed framework aims to treatment efficiency by reducing overtreatment of healthy tissue, minimising untreated regions, and shortening treatment times. The proposed system will facilitate experiments using lesion sensitive tissue-mimicking phantoms, following the methodology described by Boers et al. [13]. Future work will evaluate experimentally formed lesions against the plan along with the target region coverage. Additionally, treatment efficiency will be assessed by comparing planned and actual execution times. These metrics will evaluate the accuracy and robustness of the proposed co-robotic planning framework. We will also extend the framework to three-dimensional plans, which would further enhance its clinical viability and supporting the broader adoption of USgHIFU.

REFERENCES

[1] G. ter Haar, "HIFU Tissue Ablation: Concept and Devices," in *Therapeutic Ultrasound*, J.-M. Escoffre and A. Bouakaz, Eds. Cham: Springer International Publishing, 2016, pp. 3–20.

[2] Z. Izadifar, Z. Izadifar, D. Chapman, and P. Babyn, "An Introduction to High Intensity Focused Ultrasound: Systematic Review on Principles, Devices, and Clinical Applications," *Journal of Clinical Medicine*, vol. 9, no. 2, p. 460, Feb. 2020, number: 2.

[3] P. Gupta and A. Srivastava, "Numerical Study on the Possible Scanning Pathways to Optimize Thermal Impacts During Multiple Sonication of HIFU," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 7, pp. 2117–2128, Jul. 2021, conference Name: IEEE Transactions on Biomedical Engineering.

[4] B. Turner and D. Cranston, "A Review of High-Intensity Focused Ultrasound," *International Journal of Translational Medicine*, vol. 4, no. 1, pp. 197–207, Mar. 2024, number: 1.

[5] A. Gunderman, R. Montayre, A. Ranjan, and Y. Chen, "Review of Robot-Assisted HIFU Therapy," *Sensors (Basel, Switzerland)*, vol. 23, no. 7, p. 3707, Apr. 2023.

[6] A. Payne, U. Vyas, A. Blankespoor, D. Christensen, and R. Roemer, "Minimisation of HIFU pulse heating and interpulse cooling times," *International Journal of Hyperthermia*, vol. 26, no. 2, pp. 198–208, 2010.

[7] Y. Zhou, S. G. Kargl, and J. H. Hwang, "The Effect of the Scanning Pathway in High-Intensity Focused Ultrasound Therapy on Lesion Production," *Ultrasound in Medicine and Biology*, vol. 37, no. 9, pp. 1457–1468, 2011.

[8] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of Biomedical Optics*, vol. 15, no. 2, p. 021314, Mar. 2010.

[9] L. Howell, M. Waterston, T. M. Wah, J. H. Chandler, and J. R. McLaughlan, "A Unified Platform and Quality Assurance Framework for 3D Ultrasound Reconstruction with Robotic, Optical, and Electromagnetic Tracking," Mar. 2026, arXiv:2603.20077 [cs].

[10] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, "PLUS: open-source toolkit for ultrasound-guided intervention systems," *IEEE transactions on bio-medical engineering*, vol. 61, no. 10, pp. 2527–2537, Oct. 2014.

[11] T. Ungi, A. Lasso, and G. Fichtinger, "Open-source platforms for navigated image-guided interventions," *Medical Image Analysis*, vol. 33, pp. 181–186, Oct. 2016.

[12] "moveit/moveit2," May 2026, original-date: 2019-02-15T16:17:11Z. [Online]. Available: https://github.com/moveit/moveit2

[13] T. Boers, W. Brink, L. Bianchi, P. Saccomandi, J. van Hespen, G. Wennemars, S. Braak, M. Versluis, and S. Manohar, "An anthropomorphic thyroid phantom for ultrasound-guided radiofrequency ablation of nodules," *Medical Physics*, vol. 51, no. 2, pp. 826–838, 2024.

Dense Estimation of Retinal Tectonics via Radial Basis Function Interpolation

Tommaso Delucchi[1], Nicola Piccinelli[1], Andrea Roberti[1], Riccardo Mondin[2], Marco Pellegrini[2], and Riccardo Muradore[1]

[1]*Department of Engineering for Innovation Medicine, University of Verona, Italy*
[2]*Department of Translational Medicine, University of Ferrara, Italy*

INTRODUCTION

Estimating the retinal tectonics is crucial for understanding the structural and biomechanical changes that follow vitreoretinal surgeries. Recent studies have attempted to quantify this "en face" displacement by comparing pre- and postoperative near-infrared images. In [**?**], the retinal deformation was estimated by computing optical flow across the entire image. However, applying optical flow indiscriminately across the whole image presents significant limitations; it can inadvertently track the movement of non-retinal structures, such as choroidal vessels, or erroneously interpret common details, such as glares and local luminosity artefacts, as actual retinal movement.

To address these inaccuracies, [1] proposed to first segment retinal vessels and use them as a proxy for retinal tissue. This is biomechanically sound, as the vascular network is a primary determinant of retinal stiffness. By applying the optical flow algorithm exclusively to this network, the method isolates true retinal displacement from background noise and avoids the depth ambiguity inherent in whole-image analysis. Nevertheless, this technique inherently results in incomplete spatial coverage. Because it relies entirely on vessel presence, it cannot directly measure displacement in critical avascular areas, such as the fovea, leaving substantial portions of the functionally significant retinal field unmeasured.

In this paper, we propose a novel approach that bridges the gap between these methodologies. We utilise the segmented retinal vessels as a reliable, artefact-free proxy for localised tissue movement and subsequently apply a continuous field interpolation using Radial Basis Functions (RBF) [2]. By relying on the biomechanical assumption of tissue continuity, our method provides a complete, dense estimation of the continuous displacement field across the entire retina. This allows us to overcome the sparsity of vessel-only tracking while entirely avoiding the noise and choroidal artefacts associated with whole-image optical flow analysis.

METHODOLOGY

The proposed method, shown in Figure 1, estimates displacement through three automated steps: image registration, vessel segmentation, and dense optical flow using RBF interpolation.

First, preoperative and postoperative Near-infrared (NIR) images are aligned using Scale-Invariant Feature Transform (SIFT). Key points are extracted around the optic nerve, which serves as a stable anatomical reference. A rigid transformation model, including 2D translation and rotation, is then estimated to overlap the images without introducing artificial deformation.

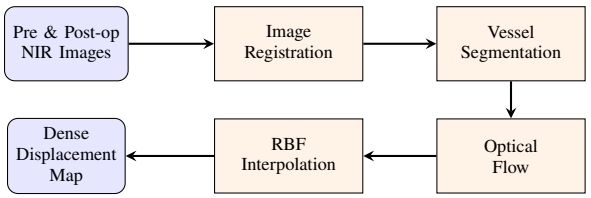


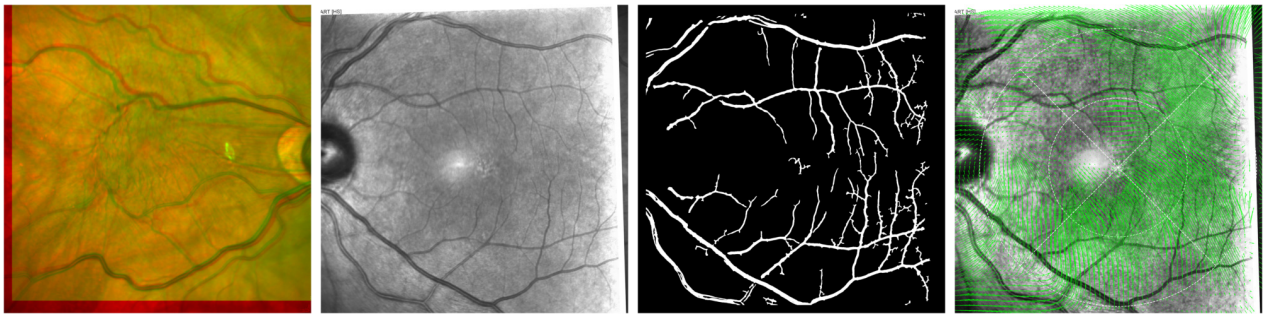
Fig. 1. Proposed retinal tectonics estimation pipeline.

Next, the retinal vascular network is segmented from the input image I and used as a reliable proxy for tissue motion. Vessels are enhanced using a multilevel Discrete Wavelet Transform (DWT) with a Daubechies-3 basis at four decomposition levels, followed by median filtering to obtain the denoised image I_W. A bilateral filter and Contrast Limited Adaptive Histogram Equalisation (CLAHE) are then applied to enhance high-frequency structures and correct uneven illumination, producing I_B. The vascular pattern is extracted using a Top-hat morphological transform and further enhanced with Hessian-based second-derivative Gaussian filters, which emphasise ridge-like tubular structures. The final binary mask is obtained by hysteresis thresholding, which preserves vessel continuity by connecting weak vessel pixels to stronger ones. Segmentation is performed in two stages to separately capture primary vessels that emerge from the optic nerve and thinner secondary branches that extend toward the macula.

Finally, Farneback optical flow is computed only within the vascular mask M, preventing the algorithm from tracking choroidal structures, reflections, or background artefacts. Farneback was preferred over local methods such as Lucas-Kanade (LK) for its ability to produce a dense field via polynomial expansion across vessel surfaces, yielding a more stable vector field $\vec{v}(x, y)$. The resulting sparse vessel-based displacement vectors $\vec{v}(x, y)$ are interpolated with Radial Basis Functions using a Thin Plate Spline kernel. Assuming retinal tissue continuity, this produces a dense displacement map across the entire macula, including avascular regions such as the fovea.

RESULTS

We evaluated our pipeline using two private patient datasets with Epiretinal Membrane (ERM) from the Hospi-



(a) Image registration (b) Preoperative NIR (c) Vessel mask (d) Dense RBF field

Fig. 2. Tectonic estimation pipeline applied to a clinical subject. (a) Qualitative assessment of the alignment between preoperative (green channel) and postoperative (red channel) NIR images. (b) Preoperative near-infrared image. (c) Segmented vascular network. (d) Final continuous displacement map reconstructed via RBF interpolation.

tal of the University of Ferrara. Dataset 1 (comprising 133 patients) was used to evaluate retinal displacement, while dataset 2 (containing 33 patients) was used to validate the registration precision. The image resolutions in the datasets are 1908×1908 and 768×768 pixels, respectively, with a spatial resolution around 4.65 µm/pixel.

To validate the automated registration module, we compared the SIFT-aligned images with those manually aligned by clinical experts (ground truth). Precision was evaluated by computing the Root Mean Square Error (RMSE) of pixel-intensity differences. For each pair of aligned images, the error is calculated as

$$RMSE = \sqrt{\frac{1}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} [I_a(x,y) - I_m(x,y)]^2} \quad (1)$$

where I_a and I_m represent the pixel intensities of the automated and manual alignments, and W and H are the image dimensions. We found an average RMSE of 8.0 units (on a 0-255 scale), showing high spatial overlap for the successfully processed subset. However, a current limitation is the reliance on keypoint density: several subjects were discarded when insufficient features were detected around the optic disc. For the retained cases, this low-intensity error confirms that the automated pipeline effectively replicates expert alignment, ensuring that the subsequent displacement analysis remains unbiased.

Our two-step filtering correctly isolated the vessels. It identified both the main vessels and the smaller branches, creating a clean mask we could safely use for optical flow computation. To measure displacement, we examined two circular areas around the fovea: an inner circle (4 mm in diameter) and an outer ring (8 mm in diameter). We split each area into four standard quadrants. As shown in Table I, our automated results are very close to the manual measurements (ground truth), with a global mean displacement of 40.83 ± 32.20 µm. As shown in Figure 2, the RBF interpolation produces a continuous, realistic map. It successfully fills the gap at the fovea without introducing the background noise typically seen when tracking the entire image.

TABLE I

RETINAL DISPLACEMENT MAGNITUDE (µm) COMPUTED ACROSS MACULAR REGIONS. COMPARISON BETWEEN THE AUTOMATED PIPELINE AND THE MANUAL GROUND TRUTH.

Anatomical Region	Quadrant	Automated Pipeline	Manual Ground Truth
Whole Macula	Global	40.83 ± 32.20	48.32 ± 36.22
Inner Ring	Average	41.88 ± 29.72	46.39 ± 32.77
Inner Ring	Superior (225)	42.91 ± 36.25	45.77 ± 34.62
Inner Ring	Inferior (45)	45.64 ± 37.77	51.20 ± 40.42
Inner Ring	East (315)	47.87 ± 42.04	50.13 ± 38.90
Inner Ring	West (135)	31.11 ± 21.70	38.46 ± 28.51
Outer Ring	Average	40.48 ± 34.12	48.96 ± 38.44
Outer Ring	Superior (225)	40.04 ± 34.01	49.84 ± 42.35
Outer Ring	Inferior (45)	41.77 ± 46.69	49.78 ± 49.50
Outer Ring	East (315)	40.96 ± 40.92	49.09 ± 38.67
Outer Ring	West (135)	38.96 ± 31.71	46.97 ± 41.36

CONCLUSION

We presented an automated pipeline for dense retinal tectonics estimation from pre- and postoperative NIR images. Combining vessel-constrained optical flow with RBF interpolation yields continuous displacement maps while mitigating artefacts from choroidal structures, reflections, and illumination. Results demonstrate good agreement with manual ground truth measurements. Currently, displacement estimation relies on the spatial distribution and orientation of retinal vessels, which can limit the detectability of certain motion directions. Future work includes broader clinical validation and improved biomechanical modelling. Extending this dense quantification to intraoperative imaging could provide valuable feedback for computer- and robot-assisted vitreoretinal surgery.

REFERENCES

[1] M. Pellegrini, G. Adamo, C. Vivarelli, A. Bizzarri, R. Mondin, P. M. Talli, F. Nasini, G. Giannaccare, and M. Mura, "Macular hole surgery and retinal tectonics: the impact of internal limiting membrane peeling size on tangential retinal displacement," *Retina*, vol. 45, no. 1, pp. 23–29, 2025.

[2] M. D. Buhmann, "Radial basis functions," *Acta numerica*, vol. 9, pp. 1–38, 2000.

[3] F. Scarinci, G. Querzoli, P. Cosimi, G. Ripandelli, M. R. Romano, A. Cacciamani, M. R. Munk, and T. Rossi, "Retinal tectonics after macular pucker surgery: Thickness changes and: En face: Displacement recovery," *Retina*, vol. 44, no. 1, pp. 102–110, 2024.

A Multimodal Laparoscopic Training Dataset for Automated Surgical Skill Assessment in Low-Resource Settings

Omar S. Choudhry[1], Sharib Ali[1], Chandra Shekhar Biyani[2], and Dominic P. Jones[3]

[1]*School of Computer Science, University of Leeds, UK*

[2]*St James's University Hospital, Leeds Teaching Hospitals NHS Trust, UK*

[3]*School of Electronic and Electrical Engineering, University of Leeds, UK*

INTRODUCTION

Minimally invasive surgery (MIS), particularly laparoscopy, reduces tissue damage, blood loss, infection and recovery time relative to open surgery. Its uptake in low- and middle-income countries (LMICs) is severely constrained by a shortage of skilled surgeons, by equipment cost, and by limited training opportunities [1], [2]. The Lancet Commission on Global Surgery estimates that 5 billion people lack access to safe surgical care and that an additional 143 million procedures are required annually [3], while the WHO records a 44.5 million shortfall of health workers, with the surgical sector disproportionately affected [4]. Expert-led skill assessment is subjective, time-consuming, and not scalable to meet the demand. Automated, vision-based skill assessment that runs on edge hardware or mobile phone devices offers a route to consistent, objective, real-time feedback during training, but it is critically bottlenecked by a lack of suitable data.

Most publicly rigid-tool laparoscopic datasets are in-vivo and image-only, e.g. the various Cholec and successive EndoVis challenge datasets, providing phase labels, tool presence or 2D boxes but no instrument pose ground truth [5], [6], [7], [8]. The few datasets which do frequently include accurate, comprehensive kinematics are robotic: JIGSAWS and ROSMA provide time-synchronised 6-DoF tip pose and jaw angle, yet captured on costly da Vinci platforms and do not represent non-robotic training or LMIC constraints [9], [10]. Training task datasets remain scarce and limited: PETRAW is entirely simulated, not representing realistic visual conditions, SimSurgSkill provides coarse VR controller data, and the WMU box-trainer provides only sparse 2D annotations without kinematics [11], [12]. Consequently, no open resource delivers continuous, ground-truth position and orientation, jaw angles plus surgeon-level skill metadata throughout a realistic non-robotic training task.

MATERIALS AND METHODS

a) Data collection: We curated the dataset across four national surgical training events: the 8th, 9th, and 10th Urology Simulation Boot Camps held in Leeds in 2023, 2024, and 2025 respectively, and the British Association of Paediatric Endoscopic Surgeons (BAPES) Annual Congress MIS course in November 2024. All trials follow the European Training in Basic Laparoscopic Urological Skills (E-BLUS) curriculum [13].

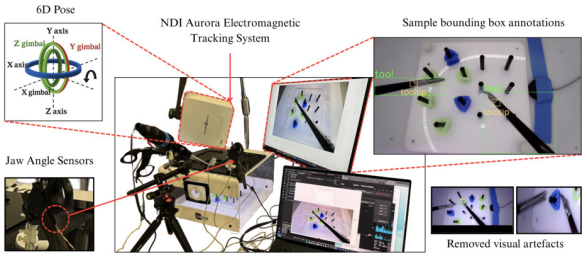


Fig. 1: Overview of the LASK data collection setup. A standard laparoscopic box trainer is augmented with four electromagnetic sensors, one per tool, one on the laparoscope, and one reference. The inset shows a sample annotated frame with tool shaft and tool tip bounding boxes and the corresponding 7-DoF kinematic stream.

The **LASK** (**LA**paroscopic **S**kill and **K**inematics) dataset captures synchronised HD endoscopic video at 30 FPS and instrument kinematics from an NDI Aurora electromagnetic tracker [14] (Figure 1). Four sensors are used: one on each tool shaft, one on the laparoscope, and one stationary world reference. Each tool sensor is then calibrated to its tool tip in three stages: (i) rotating the tool about its shaft axis to recover the sensor offset relative to the central rotation axis; (ii) attaching a pointed cap to the end of the tool at a known offset from the jaw hinge and running a pivot calibration to register that point; and (iii) fixing the shaft flat against the world reference to recover the axial rotation around the shaft. Position and unit-quaternion orientation are then reported at the tool tip, yielding 6-DoF pose. All trials were recorded on a non-ferrous box trainer to keep metallic interference well below the system's published drift threshold; the rigid tool shafts and light peg-transfer loads produced no measurable bending, so the coil-to-tooltip transformation remained fixed across each trial. From the 9th Boot Camp onwards, magnetic jaw aperture sensors on the tool handles add a seventh degree of freedom, smoothed using a second-order Butterworth low-pass filter

b) Participant survey: At each event participants completed a short pre-trial questionnaire covering (i) the Edinburgh Handedness Inventory [15]; (ii) lifetime laparoscopic case count across the surgeon's career; and (iii)

laparoscopic case count over the last 12 months as a recency-of-practice indicator. Skill bands are loosely defined by lifetime cases: novice (< 20 cases), intermediate (20–100 cases), and expert (> 100 cases).

c) Annotation: On the peg transfer subset of LASK, every 100$^{\text{th}}$ frame is annotated with per-tool segmentation polygons and 8 keypoints (4 on the shaft, joint, tip, and 2 jaw tips) across 38 trials, yielding 882 fully-completed positive frames after quality control to remove over-lit, cut-off, or corrupted images. Following the convention in [16], frames in which neither tool is in scene are kept as labelled negatives, providing a further 8,226 background samples that materially reduce false positives in downstream detectors. A separate, earlier set of annotations produced 3,725 bounding-box labels used in our detection benchmark [17]. Finally, a new dual-tool phase layer was added in this work: each annotated frame is labelled with one of seven phases (idle, reach, nudge, grasp, transfer, place, dropped) per tool.

RESULTS

TABLE I: The LASK dataset at a glance.

Descriptor	Dataset Statistics
Trials (peg transfer / circle cutting / knot tying)	114 / 21 / 9
Unique surgeons	144
novice (< 20 cases)	56
intermediate (20–100 cases)	42
expert (> 100 cases)	46
Adult urology / paediatric surgeons	114 / 30
Handedness, right / left / mixed	133 / 5 / 6
Total recorded duration	≈4.3 h
Frames at 30 FPS	469,904
Peg transfer annotations	
Tool segmentation + 8 keypoints (sampled)	882 frames, 38 trials
Labelled negative (no tool in scene) samples	8,226 frames
Bounding-box annotations [17]	3,725 frames
Dual tool phase and cycle labels	96,083 frames, 457 cycles

LASK currently comprises 469,904 video frames across 144 trials from 144 unique surgeons. It is the first non-robotic peg transfer dataset that combines continuous 7-DoF instrument kinematics, jaw aperture sensing, surgeon-level skill metadata, dual tool phase and cycle labels, and a documented out-of-distribution evaluation split (Cohort C, n=24). Building on this dataset, we have already published a benchmarking study of state-of-the-art object detectors for laparoscopic tool detection in low-resource settings [17], and developed a probabilistic 7-DoF vision-only tool pose tracker that achieves 6.8 mm position and 11.9° rotation RMSE with calibrated per-frame uncertainty [18].

CONCLUSIONS AND DISCUSSION

We introduce the novel dataset LASK, enabling researchers to develop end-to-end automated laparoscopic skill assessment pipelines that are realistic for low-resource training. The combination of continuous kinematic ground truth, jaw-aperture sensing, and dual-tool phase annotation on a non-robotic peg transfer addresses three independent gaps in existing public datasets: the absence of marker-less ground-truth kinematics, the robotic-only nature of kinematics-rich datasets and the lack of granularity in phase annotations for the peg transfer task in real training data. Future work will use the new phase and cycle annotations to train a gesture-recognition model utilising the tool, peg, and pegboard segmentation and keypoints.

REFERENCES

[1] T. A. Rockall and K. Sasapu, "Laparoscopy in the developing world: An advanced technique to improve global surgical care?" *International Journal of Surgery*, vol. 12, pp. S203–S207, 2014.

[2] S. Bodenstedt, M. Allan, A. Agustinos, X. Du, L. Garcia-Peraza-Herrera, H. Kenngott, T. Kurmann *et al.*, "Comparative evaluation of instrument segmentation and tracking methods in minimally invasive surgery," *arXiv preprint arXiv:1805.02475*, 2018.

[3] J. G. Meara, A. J. M. Leather, L. Hagander, B. C. Alkire, N. Alonso, E. A. Ameh, S. W. Bickler *et al.*, "Global surgery 2030: evidence and solutions for achieving health, welfare, and economic development," *The Lancet*, vol. 386, no. 9993, pp. 569–624, 2015.

[4] World Health Organization, "World health statistics 2016: Monitoring health for the sdgs, sustainable development goals," https://www.who.int/data/gho/publications/world-health-statistics, Geneva, 2016.

[5] S. Ali *et al.*, "A comprehensive survey on recent deep-learning approaches for laparoscopic instrument detection, segmentation and tracking," *arXiv preprint arXiv:2304.11150*, 2023.

[6] C. I. Nwoye, et al., "CholecTrack20: A dataset for multi-class multiple tool tracking in laparoscopic surgery," 2023, version Number: 1. [Online]. Available: https://arxiv.org/abs/2312.07352

[7] W.-Y. Hong, C.-L. Kao, Y.-H. Kuo, J.-R. Wang, W.-L. Chang, and C.-S. Shih, "CholecSeg8k: A Semantic Segmentation Dataset for Laparoscopic Cholecystectomy Based on Cholec80," Dec. 2020, arXiv:2012.12453 [cs]. [Online]. Available: http://arxiv.org/abs/2012.12453

[8] (2015) EndoVis Sub-Challenge Instrument Segmentation and Tracking - grand challenge. grand-challenge.org. Https://endovissub-instrument.grand-challenge.org/Data/. [Online]. Available: https://endovissub-instrument.grand-challenge.org/Data/

[9] Y. Gao, S. S. Vedula, C. E. Reiley, N. Ahmidi, B. Varadarajan, H. C. Lin, L. Tao, L. Zappella, B. Bejar, D. D. Yuh, C. C. G. Chen, R. Vidal, S. Khudanpur, and G. D. Hager, "JHU-ISI Gesture and Skill Assessment Working Set (JIGSAWS): A Surgical Activity Dataset for Human Motion Modeling," 2015.

[10] I. Rivas-Blanco, C. J. Perez-del Pulgar, A. Mariani, C. Quaglia, G. Tortora, V. F. Munoz, and A. Menciassi, "Surgical robotics: A ROS-Based open-source modular simulator for the da Vinci robot — the rosma dataset," *Sensors*, vol. 23, no. 1, p. 302, 2023.

[11] A. Huaulmé, D. Sarikaya, K. Le Mut, F. Despinoy, Y. Long, Q. Dou *et al.*, "PETRAW: A peg-transfer workflow recognition challenge with multi-modal data," *Computer Methods and Programs in Biomedicine*, vol. 234, p. 107533, 2023.

[12] F. R. Fathabadi, J. L. Grantner, S. A. Shebrain, and I. Abdel-Qader, "Surgical skill training and evaluation for a peg transfer task of a three-camera laparoscopic box trainer system," *Sensors*, vol. 22, no. 21, p. 8261, 2022.

[13] European Association of Urology, "E-BLUS examination: European training in basic laparoscopic urological skills," https://uroweb.org/education-careers/esu/courses/eblus, 2016.

[14] J. B. Hummel, M. R. Bax, M. L. Figl, Y. Kang, C. Maurer, W. W. Birkfellner, H. Bergmann, and R. Shahidi, "Design and application of an assessment protocol for electromagnetic tracking systems," *Medical Physics*, vol. 32, no. 7Part1, pp. 2371–2379, 2005.

[15] R. C. Oldfield, "The assessment and analysis of handedness: The edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.

[16] S. M. K. Hasan and C. A. Linte, "Detection, segmentation, and 3D pose estimation of surgical tools using convolutional neural networks and algebraic geometry," *Medical Image Analysis*, vol. 70, p. 101994, 2021.

[17] O. S. Choudhry, S. Ali, C. S. Biyani, and D. P. Jones, "Real-time tool detection in laparoscopic datasets for surgical training in low-resource settings," in *Joint 17th AE-CAI / CARE / OR 2.0 Workshop at MICCAI 2025*, 2025, special issue, IET Healthcare Technology Letters.

[18] O. S. Choudhry, S. Ali, C. S. Biyani, and D. P. Jones, "Bayesian temporal pose networks for uncertainty-calibrated laparoscopic tool pose tracking," 2026, MICCAI 2026 submission; under review.

Benchmarking object detectors for naso-endoscopic anatomical features

M. M. Loureiro da Rocha[1,2], W. Jansen[1], C. D. Rocha[3], A. J. Sousa[3,4], M. J. A. van Alphen[2], M. W. M. van den Brekel[2], and F.J. Siepel[1]

[1]*Robotics and Mechatronics Group, University of Twente, the Netherlands*
[2]*Department of Head and Neck Oncology, the Netherlands Cancer Institute, the Netherlands*
[3]Institute for Systems and Computer Engineering, Technology and Science, Porto, Portugal
[4]Faculdade de Engenharia, Universidade do Porto, Porto, Portugal

INTRODUCTION

Flexible transnasal endoscopy (TE) is a gold-standard diagnostic tool for many upper gastrointestinal and airway-related symptoms, making it one of the most frequently performed procedures in clinical practice [1]. However, the growing demand for endoscopic procedures, combined with global shortages of trained healthcare personnel, has made the availability of qualified clinicians a limiting factor. As a result, clinicians are often required to perform higher procedure volumes [2]. TE is typically performed by dedicated specialists who require substantial training [3]. This requirement restricts the expansion of endoscopy practice to other healthcare providers such as nurses or general practitioners. The impact is particularly critical in remote or underserved regions, where access to qualified personnel may require several hours of travel. The aforementioned challenges highlight the need for improved training methods and procedural support systems. The integration of artificial intelligence (AI) into endoscopic training workflows could support skill acquisition by providing feedback and guidance during training. This may lower the barrier to proficiency for early-career clinicians and enable a broader group of healthcare providers to safely perform endoscopic procedures. Such systems can be embedded into robotic platforms with haptic feedback to improve guidance, ergonomics and procedural stability [4]. By suggesting optimized movements, they may assist clinicians during navigation, which could reduce unintended contact with sensitive anatomical regions and improve patient comfort [5]. In this work, we present a benchmarking study of multiple real-time object detection networks for identifying key anatomical structures (orientation anchors) during TE examinations. Reliable detection of these structures is a fundamental requirement for AI-assisted training and navigation systems. The evaluated detectors therefore establish a technical foundation towards the development of intelligent support and robotic navigation frameworks for endoscopic procedures.

MATERIALS AND METHODS

We fine-tuned seven different object detector architectures pretrained on the COCO dataset [6]: YOLO11-n/s, YOLO12-n/s, YOLO26-n/s, and RT-DETR4-s. The models were trained to detect the epiglottis, the airway and the vocal folds on a dataset composed of 18260 frames from 13 TE videos of different patients, that contained 8,499 vocal fold, 10,351 airway, and 10,390 epiglottis instances, manually annotated using CVAT (Fig. 1).

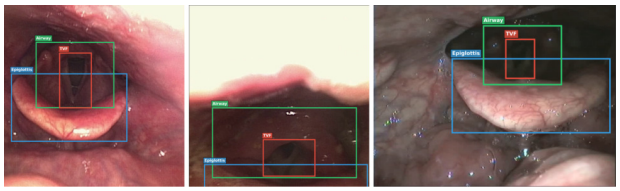


Fig. 1: Annotation examples of the three classes: airway (green), vocal folds (TVF, red) and epiglottis (blue).

To ensure a controlled and fair benchmarking process, training was performed on the same NVIDIA L40 48 GB GPU, and identical training hyperparameters were used across all model architectures. The number of epochs and early-stopping patience were determined through a preliminary 5-fold cross-validation on the YOLO11-n model. All remaining hyperparameters were kept at their default optimized settings for the COCO dataset [6]. The complete set of hyperparameters is listed in Table I.

TABLE I: Standardized training hyperparameters for model benchmarking.

Hyperparameter	**Value**
Epochs	300
Patience	50
Batch size	16
Image size	640
Initial learning rate	0.001
Final learning rate factor	0.01
Optimizer	AdamW/muSGD
Weight decay	0.0005

Due to the limited dataset size, a 13-fold leave-one-patient-out cross-validation scheme was employed, in which one patient was reserved for testing in each fold, and the remaining 12 patients were used for training (10) and validation (2). This procedure resulted in 13 models per architecture, and the final performance was obtained by averaging across all folds. The evaluation metrics included *Precision* (P), *Recall* (R), *F1-score* (F1),

mean average precision (AP@50, AP@50-95) and scale-normalized centroid error ($\tilde{d}_c$). All metrics were computed for a confidence of 0.25 and an intersection-over-union threshold of 0.5. To improve model generalization, the training and validation sets underwent random data augmentation, including scaling, translation, mosaic, random erasing, vertical flipping, and brightness/contrast adjustments.

RESULTS

Table II summarizes the average detection results. When evaluated using the *F1-score*, RT-DETR4-s achieved the highest overall performance, with an overall *F1-score* of 0.87 and a *Recall* of 0.92. The highest *Precision* among all models was observed for YOLO26-s (0.90), while the best detection accuracy in terms of AP50 and AP50-95 was obtained by YOLO12-s (0.95 and 0.48, respectively). The lowest centroid error ($\tilde{d}_c$) was achieved by RT-DETR4-s (0.03), indicating improved spatial consistency compared to the YOLO-based models, which showed similar error levels around 0.04. The highest *F1-score* for the epiglottis and airway classes was obtained by RT-DETR4-s (0.89 and 0.87, respectively). For the vocal folds class comparable performance was observed across YOLO11-s, YOLO12-n, YOLO26-s and RT-DETR4-s. Across all evaluated models, the epiglottis class achieved the highest detection performance compared to the airway and vocal folds classes.

TABLE II: Average performance metrics across all folds. Bold indicates the best overall performance per metric.

Model	Class	P	R	F1	AP50	AP50-95	$\tilde{d}_c$
YOLO11-n	Epiglottis	0.90	0.83	0.86	0.98	0.58	0.03
	Airway	0.84	0.80	0.82	0.94	0.34	0.05
	Vocal Folds	0.86	0.81	0.83	0.90	0.40	0.05
	Overall	0.86	0.81	0.84	0.94	0.44	0.04
YOLO11-s	Epiglottis	0.88	0.83	0.85	0.97	0.59	0.03
	Airway	0.86	0.81	0.84	0.95	0.37	0.05
	Vocal Folds	0.90	0.80	0.85	0.88	0.44	0.05
	Overall	0.88	0.81	0.85	0.94	0.47	0.04
YOLO12-n	Epiglottis	0.91	0.80	0.85	0.98	0.62	0.03
	Airway	0.87	0.76	0.82	0.91	0.29	0.05
	Vocal Folds	0.90	0.81	0.85	0.84	0.37	0.05
	Overall	0.89	0.79	0.84	0.91	0.43	0.04
YOLO12-s	Epiglottis	0.92	0.79	0.85	0.99	0.64	0.03
	Airway	0.87	0.78	0.82	0.93	0.33	0.05
	Vocal Folds	0.88	0.80	0.84	0.94	0.48	0.05
	Overall	0.89	0.79	0.84	**0.95**	**0.48**	0.04
YOLO26-n	Epiglottis	0.92	0.79	0.85	0.98	0.60	0.03
	Airway	0.86	0.77	0.81	0.85	0.26	0.05
	Vocal Folds	0.91	0.79	0.84	0.91	0.46	0.05
	Overall	0.89	0.78	0.83	0.91	0.44	0.04
YOLO26-s	Epiglottis	0.92	0.79	0.85	0.99	0.62	0.02
	Airway	0.86	0.74	0.80	0.88	0.33	0.05
	Vocal Folds	0.91	0.79	0.85	0.89	0.48	0.05
	Overall	**0.90**	0.77	0.83	0.92	0.47	0.04
RT-DETR4-s	Epiglottis	0.85	0.94	0.89	0.96	0.60	0.02
	Airway	0.83	0.91	0.87	0.80	0.26	0.04
	Vocal Folds	0.80	0.90	0.85	0.88	0.43	0.04
	Overall	0.83	**0.92**	**0.87**	0.88	0.43	**0.03**

DISCUSSION AND CONCLUSIONS

The performance variation between architectures and classes suggests that different architectures provide complementary strengths. While YOLO-based models achieved strong detection accuracy, RT-DETR4-s demonstrated superior *Recall* and localization performance. This complementary behaviour indicates that the design of ensemble predictors combining multiple detectors may improve robustness across anatomical structures. Performance may also benefit from the incorporation of anatomical priors during training, for example through geometric loss functions that encourage anatomically consistent predictions. Additionally, integrating temporal smoothing methods, such as Kalman filtering, could improve the stability and continuity of predictions across consecutive frames. The higher detection performance observed for the epiglottis class is likely related to its visual characteristics. The epiglottis typically appears in high-reflection regions (cf. Figure 1), whereas the airway and vocal folds are often located in lower-light areas, making them more difficult to distinguish.

Future work should extend the training framework beyond the laryngeal structures considered in this study to include additional anatomical regions. In particular, regions of the nasopharynx and oropharynx may provide useful anatomical references for navigation and structural identification. Expanding the system toward applications supporting cancer screening may also enable earlier detection of pathological findings. Within the broader trajectory toward robotic system integration, this benchmarking study represents an initial step toward a more comprehensive framework capable of navigating from the nasal cavity to the esophagus. Such a system could incorporate visual SLAM (Simultaneous Localization and Mapping) algorithms to provide spatial awareness and support procedural guidance during training/examination. Overall, this work provides a comparative evaluation of object detection architectures for identifying key anatomical structures during TE. The low scale-normalized centroid errors demonstrate that these models can accurately localize anatomical landmarks, supporting their use in future guidance systems.

REFERENCES

[1] N. Nguyen, J. Mark, and G. T. Furuta, "Emerging role of transnasal endoscopy in children and adults," *Clinical Gastroenterology and Hepatology*, vol. 20, no. 3, pp. 501–504, 2022.

[2] H. Hayes, N. Lenze, H. J. Ritz, V. Pandian, M. J. Brenner, and P. Yalamanchi, "Workforce trends and evolving roles of physician assistants in otolaryngology: A national longitudinal survey," *Annals of Otology, Rhinology & Laryngology*, p. 00034894251409704, 2026.

[3] E. M. Benadom and N. L. Potter, "The use of simulation in training graduate students to perform transnasal endoscopy," *Dysphagia*, vol. 26, no. 4, pp. 352–360, 2011.

[4] E. L. Servais, L. Rashidi, P. Porwal, M. Garibaldi, and A. J. Hung, "Novel force feedback technology improves suturing in robotic-assisted surgery: a pre-clinical study," *Surgical Endoscopy*, vol. 39, no. 2, pp. 1217–1226, 2025.

[5] S. Alvi and P. Harsha, "Flexible nasopharyngoscopy," in *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2026, [Updated 2023 Aug 8]. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK539740/

[6] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.

A Pre-Experimental Scan-Planning Simulator for Hexapod-Mounted On-Chip Terahertz Sensing

Nikita J. Greenidge, Connor D. W. Mosley, Josh P. R. Nixon, John E. Cunningham, Pietro Valdastri

School of Electronic and Electrical Engineering, University of Leeds, Leeds, UK

n.j.greenidge@leeds.ac.uk

INTRODUCTION

Surgical margin assessment requires sub-millimetre tissue characterisation. Terahertz (THz) sensing offers a non-ionising, label-free contrast mechanism [1], [2], [3], and on-chip quarter-wave stub (QWS) resonators convert local sample loading into a measurable resonant-frequency shift [4], [5], [6]. Robotic platforms can raster the sample to form a spatial map [7], but the experimental outcome depends on interacting parameters: sensor design (stub length, substrate index), chip–sample separation, robot motion (retract height, acceleration profile), dwell protocol (settle time, pulse count, acquisition rate), and raster strategy (pitch, pattern, extent). Iterating over these empirically with a robotic THz platform is time-prohibitive.

Here, scan planning is treated as a robot-constrained sensing problem: select sensor, motion, dwell, and raster parameters to satisfy a spatial-resolution target while minimising predicted execution time. We present a Python simulator that couples all of these parameter groups, enabling pre-experimental planning: any subset of parameters can be swept in silico, the predicted $f_{\mathrm{res}}(x, y)$ map and scan time computed for each combination, and the best configuration chosen before committing hardware time. In this paper we apply the simulator to a silicone bowtie phantom, present a raster-pitch sweep as the primary illustrative case, and compare predictions against a measured raster.

MATERIALS AND METHODS

Robotic THz platform. In this work we use a Symétrie SOLANO hexapod rastering the sample beneath a fixed on-chip QWS band-stop resonator [4], as shown in Fig. 1(A).

On-chip sensor model. Simulated pulses follow [6]: a sech^2 input is Fourier transformed, a 15 dB Lorentzian notch is applied at

$$f_0 = \frac{c}{4\,l\,\sqrt{\varepsilon_{\mathrm{eff}}}}, \tag{1}$$

where c and l are the speed of light and stub length, and $\varepsilon_{\mathrm{eff}} = \big((n_{\mathrm{sample}} + n_{\mathrm{sub}})/2\big)^2$ is the effective permittivity, taken as the square of the arithmetic mean of the sample ($n_{\mathrm{sample}} = 1.41$ for Ecoflex 30) and quartz substrate ($n_{\mathrm{sub}} = 2.09$) refractive indices; the output is inverse Fourier transformed and fit with a sloped Lorentzian to extract f_{res} (Fig. 1(B,C)). Chip–sample separation is modelled by fitting Ansys HFSS finite-element simulations of resonant frequency vs. gap to

$$f(d) = \left(\frac{\alpha\, d}{d + d_0} + \beta\right) f_{\mathrm{air}}, \tag{2}$$

where d_0 is the characteristic gap scale, f_{air} is the air-loaded resonance, and $\alpha + \beta = 1$. The fitted parameters are specific to this chip–sample pair (Ecoflex 30, $n = 1.41$, embedded in the HFSS model): $(\alpha,\, d_0,\, \beta) = (0.11,\, 5\,\mu\mathrm{m},\, 0.89)$, $f_{\mathrm{air}} = 233.5$ GHz (Fig. 1(D)).

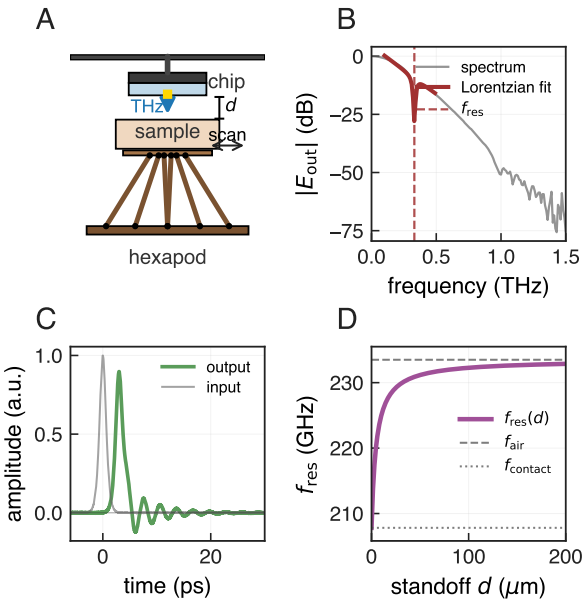


Fig. 1: Robotic THz platform and on-chip sensor model. (A) The SOLANO hexapod translates the sample beneath a fixed on-chip QWS sensor with local standoff d. (B) Simulated output spectrum with sloped Lorentzian fit used to extract f_{res} (Ecoflex 30, $n = 1.41$). (C) Input and output time-domain waveforms. (D) Calibrated chip–sample separation response.

Robot motion model. Scan time is computed from a trapezoidal velocity profile applied to each axis ($v^{xy}_{\max} = 30$ mm/s, $v^{z}_{\max} = 20$ mm/s, SOLANO datasheet; maximum operational acceleration $a_{\max} = 200$ mm/s^2). Each pixel follows a touchdown cycle: retract, translate XY, lower to contact, settle, and integrate N_p pulses at rate f_p.

Parameter sweep. We sweep the motion and raster parameter groups. For the motion group, at the 0.5 mm

Z retract and all tested XY pitches both axes operate in the triangular regime ($d < v_{\max}^2/a_{\max}$), so peak speed is never reached and scan time is independent of $v_{\max}$; the Z-cycle (lower and raise, 200 ms) dominates the 320 ms per-pixel overhead, making retract height the principal time lever after pitch. For the raster group, pitches from 20 to 200 µm are swept against a converged 1 µm reference; the trade-off between scan time and spatial resolution (RMSE) identifies 50 µm as the point of diminishing returns, with coarser pitches degrading boundary-position accuracy ($\Delta/2 = 25\,\mu\text{m}$ at 50 µm) (Figs. 2(D–F,G)).

Experimental validation. Using the 50 µm raster pitch selected from the parameter sweep, we scanned a 1.6 mm square region of an Ecoflex 30 silicone bowtie phantom ($W = H = 1.40$ mm, waist 0.46 mm, measured THz refractive index $n = 1.41$; edge taper 200 µm ramping standoff from contact at the centre to 50 µm at the edge), replicated in simulation. At each grid point the stage settled 20 ms; 100 THz waveforms were acquired at 1 kHz, averaged, and fit with a sloped Lorentzian to extract f_{res}; no pixels were discarded. Predictions are compared against the measured raster by rigid registration (translation and rotation minimising pixel-wise squared error); RMSE serves as the comparison metric (Figs. 2(A–C)).

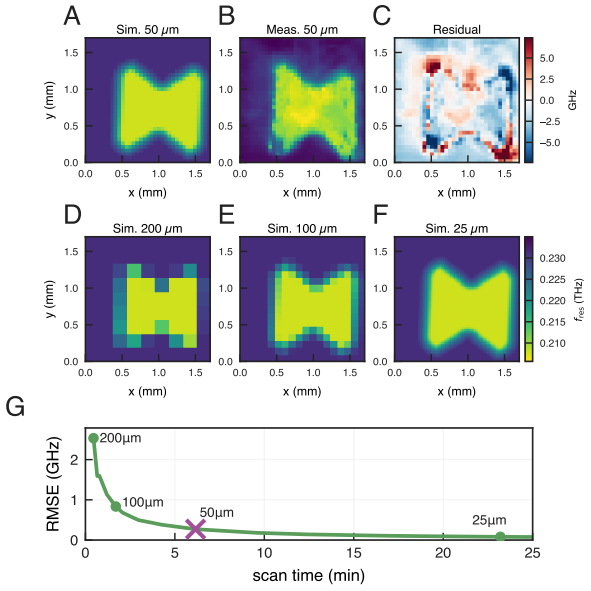


Fig. 2: Parameter sweep and experimental comparison. (A) Simulated f_{res} map at 50 µm pitch. (B) Measured 1.6 mm square raster at 50 µm pitch. (C) Pixel-wise residual, simulated minus measured, after rigid registration; colour scale clipped to ±5 GHz; peak 16.30 GHz. (D–F) Simulated f_{res} maps at 200 µm, 100 µm, and 25 µm pitch, illustrating the scan-time–spatial-resolution trade-off. (G) Simulated RMSE vs. predicted scan time; magenta cross marks the selected 50 µm pitch.

CONCLUSIONS AND DISCUSSION

The motion sweep reveals that retract is the single most time-consuming parameter: at the sub-millimetre distances involved, both Z and XY axes operate in the triangular kinematic regime where $v_{\max}$ is never reached, and the Z-cycle (lower and raise, 200 ms) accounts for 63% of the per-pixel overhead. After retract, pitch is the principal remaining time lever; the raster sweep identified 50 µm as the best balance of scan time and spatial resolution — coarser pitches degrade boundary localisation without recovering scan time, while finer pitches yield diminishing gains in spatial resolution (RMSE).

The measured comparison shows that the simulator reproduces the gross bowtie footprint, but the measured RMSE 2.82 GHz ($\sigma = 2.69$ GHz) exceeds the pitch-limited discretisation error (0.27 GHz) by an order of magnitude. Comparing simulation to experiment carries inherent residual error: phantom alignment uncertainty and physical effects not captured by the model — surface roughness and localised air gaps — introduce contact variation that cannot be predicted from geometry alone. Incorporating contact or force measurement during scanning would allow these variations to be detected and, in future closed-loop operation, potentially reduced.

The simulator is parameterised by sensor calibration and kinematic constraints, allowing it to be re-targeted to other on-chip QWS designs and robotic platforms when the corresponding calibration and motion parameters are available. Future work will explore ex vivo tissue and more clinical use cases, as well as integrating a force sensor into the experimental setup for closed-loop standoff control.

REFERENCES

[1] A. J. Fitzgerald, V. P. Wallace, M. Jimenez-Linan, L. Bobrow, R. J. Pye, A. D. Purushotham, and D. D. Arnone, "Terahertz pulsed imaging of human breast tumors," *Radiology*, vol. 239, no. 2, pp. 533–540, May 2006.

[2] P. C. Ashworth, E. Pickwell-MacPherson, E. Provenzano, S. E. Pinder, A. D. Purushotham, M. Pepper, and V. P. Wallace, "Terahertz pulsed spectroscopy of freshly excised human breast cancer," *Optics Express*, vol. 17, no. 15, pp. 12 444–12 454, 2009.

[3] J. J. Young, B. G. Page, A. I. Hernandez-Serrano, J. Hardwicke, and E. Pickwell-MacPherson, "In vivo terahertz sensing of skin cancer patients," *Advanced Photonics Research*, vol. 6, no. 6, p. 2400195, 2025.

[4] L. Dazhang, J. Cunningham, M. B. Byrne, S. Khanna, C. D. Wood, A. D. Burnett, S. M. Ershad, E. H. Linfield, and A. G. Davies, "On-chip terahertz Goubau-line waveguides with integrated photoconductive emitters and mode-discriminating detectors," *Applied Physics Letters*, vol. 95, no. 9, p. 092903, 2009.

[5] S. J. Park, R. Tucker, E. Pickwell-MacPherson, and J. E. Cunningham, "Design of a split ring resonator integrated with on-chip terahertz waveguides for colon cancer detection," *Advanced Theory and Simulations*, vol. 5, no. 9, p. 2200313, 2022.

[6] J. P. R. Nixon, C. D. W. Mosley, S. J. Park, C. D. Wood, and J. Cunningham, "Simulated comparison of on-chip terahertz filters for sub-wavelength dielectric sensing," *Sensors*, vol. 26, no. 1, p. 129, Jan. 2026.

[7] A. Dogra, D. Jones, A. I. Hernandez Serrano, S. Chakraborty, J. J. Young, B. G. Page, J. Hardwicke, P. Valdastri, and E. Pickwell-MacPherson, "Towards autonomous robotic THz-based in vivo skin sensing: the PicoBot," *Scientific Reports*, vol. 15, no. 1, p. 4568, Feb. 2025.

Trackerless Freehand Ultrasound 3D Reconstruction in Orthopedic Surgery

Jinghao Fei, Ruixuan Li, Carlos Rodriguez-Guerrero, and Emmanuel Vander Poorten

Robot Assisted Surgery Group, KU Leuven, Belgium

INTRODUCTION

In orthopedic surgery, intraoperative imaging is critical for surgical planning and navigation. While computed tomography (CT) and fluoroscopy are traditionally used, they expose patients and surgeons to ionizing radiation. Ultrasound (US) imaging presents a compelling, radiation-free alternative due to its non-invasive nature, real-time feedback, and cost-effectiveness. However, traditional 2D US lacks spatial context. Translating this modality into 3D requires precise spatial coordinates for each image frame. Traditional solutions rely on external tracking devices (e.g., optical trackers or robotic arms), which clutter the operating room, complicate workflow, and suffer from line-of-sight occlusion.

To address these limitations, researchers investigated trackerless approaches. Existing methods estimate probe motion directly from 2D images using speckle decorrelation [1] or deep learning architectures that directly regress frame-to-frame spatial transformations from raw images [2]. However, these image-based methods struggle in low-texture tissues and are sensitive to complex scanning trajectories.

This work introduces a multi-sensor fusion approach relying on low-cost on-board inertial and tactile sensors to estimate the US probe's trajectory and produce a 3D US reconstruction.

METHODS

The system integrates a US probe (Fujifilm SonoSite M-Turbo®), an inertial measurement unit (IMU, LSM6DS3TR-C, STMicroelectronics®) and a tactile sensor (IP-S1610-Elite PXSR-STDIP03B, PaXini Tech). Probe pose estimation is formulated as a sequential prediction task. Given N consecutive blocks, each consisting of K stacked US frames, the proposed multimodal temporal fusion network predicts the $N-1$ relative transformations between adjacent blocks.

The network contains three parallel feature extraction branches, as shown in Fig. 1. A shared-weight EfficientNet-B1 [3] extracts spatial features from each US frame, which are then fed into a two-layer unidirectional Gated Recurrent Unit (GRU) to capture long-term spatio-temporal dependencies, yielding image features $\mathbf{H}_{\text{img}} \in \mathbb{R}^{B\times N\times 512}$. A 1D CNN processes a temporal window of past sensor readings aligned with each US frame and, after adaptive average pooling, outputs sensor trend features $\mathbf{H}_{\text{cnn}} \in \mathbb{R}^{B\times N\times 128}$. A parallel MLP branch processes the instantaneous tactile and IMU readings, which are synchronised with the current US frame. This produces instantaneous state features $\mathbf{H}_{\text{inst}} \in \mathbb{R}^{B\times N\times 64}$. For each pair of consecutive blocks, the three feature vectors are concatenated channel-wise. A shared linear head regresses the 9-DoF relative transformation (6D rotation [4] + 3D translation). Gram–Schmidt orthogonalization is applied to recover valid SE(3) matrices.

This project received funding from the FWO project NO.12A8N26N.

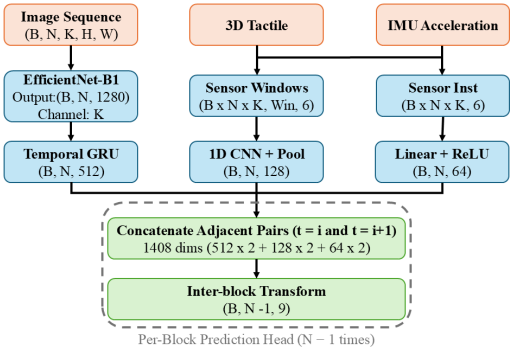


Fig. 1: Overview of the proposed network structure. B is the batch size; N is the number of blocks in a sequence; K is the number of frames in a block; H and W represent the height and width of the image; and *Win* is the sensor windows.

The network is trained with a composite loss function:

$$\mathcal{L} = \lambda_t \mathcal{L}_{\text{traj}} + \lambda_a \mathcal{L}_{\text{aux}} + \lambda_e \mathcal{L}_{\text{edge}}.$$

Trajectory accumulation loss $\mathcal{L}_{\text{traj}}$ serves as the primary supervision. At inference, predicted relative transformations are chained into absolute poses, so local errors accumulate as a random walk. To penalize this drift during training, the $N-1$ relative transformations $\Delta\mathbf{T}_i$ are chained from a ground-truth (GT) anchor $\mathbf{T}^{\text{init}}$:

$$\mathbf{T}_i^{\text{pred}} = \mathbf{T}^{\text{init}} \prod_{j=0}^{i-1} \Delta\mathbf{T}_j, \quad i = 1, 2, \ldots, N-1,$$

The accumulated poses are applied to the four calibrated corner points of the US image plane: $\mathbf{c}_{i,k}^{\text{pred}} = \mathbf{T}_i^{\text{pred}} \cdot \mathbf{T}_{\text{calib}} \cdot \mathbf{p}_k$ and $\mathcal{L}_{\text{traj}}$ is the mean squared error between predicted and GT corners over all representative frames:

$$\mathcal{L}_{\text{traj}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{1}{4} \sum_{k=0}^{3} \|\mathbf{c}_{i,k}^{\text{pred}} - \mathbf{c}_{i,k}^{\text{gt}}\|^2.$$

Auxiliary relative accuracy loss $\mathcal{L}_{\text{aux}}$ stabilizes training, since relying solely on $\mathcal{L}_{\text{traj}}$ can cause vanishing gradients through the long chain of matrix products. It compares each predicted relative transformation directly with the GT in the probe coordinate frame $\mathbf{p}_k^{\text{tool}}$:

$$\mathcal{L}_{\text{aux}} = \frac{1}{N-1} \sum_{i=0}^{N-2} \frac{1}{4} \sum_{k=0}^{3} \|\Delta\mathbf{T}_i^{\text{pred}} \cdot \mathbf{p}_k^{\text{tool}} - \Delta\mathbf{T}_i^{\text{gt}} \cdot \mathbf{p}_k^{\text{tool}}\|^2.$$

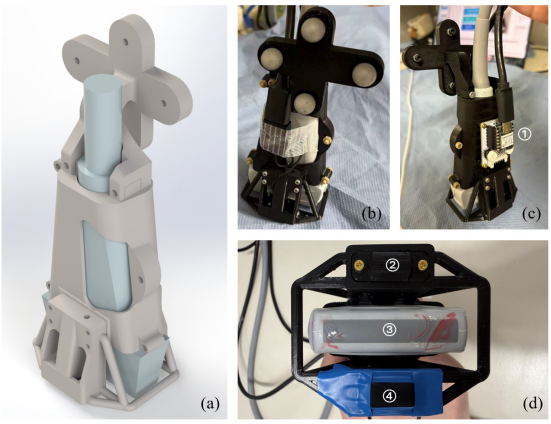


Fig. 2: (a) CAD model of the sensor-integrated US probe. (b, c, d) Front, Rear, Bottom view. ① the IMU, ② the dummy sensor, ③ the US probe, and ④ the tactile sensor.

Geometric regularisation loss $\mathcal{L}_{\text{edge}}$ prevents non-physical deformation of the predicted US plane by constraining its aspect ratios to match those of the GT. For predicted or GT edge, the four edge lengths e_j are normalized by the first edge e_o: $r_j = e_j/e_0$. The loss is expressed as:

$$\mathcal{L}_{\text{edge}} = \frac{1}{4}\sum_{j=0}^{3}\left|r_j^{\text{pred}} - r_j^{\text{gt}}\right|.$$

EXPERIMENTS AND RESULTS

The IMU (① in Figure 2(c)) is embedded within a Seeed Studio XIAO nRF52840 Sense microcontroller. On the opposite side of the US probe (③ in Figure 2(d)), a tactile sensor (④) is attached, with its sensing surface positioned 2 mm below the skin-contact plane to capture normal and shear forces. A dummy sensor (②) of identical form factor is mounted on the the IMU side to keep the probe perpendicular to the skin. GT poses were collected using an optical marker attached to the probe.

Data were collected from the forearms and lower legs of two authors, yielding 40 sequences (approximately 1360 seconds in total) of reciprocating straight, curved, and zigzag trajectories, and split as 6:2:2 into training, validation, and test sets for 3000 epochs of training. The US probe was operated at a frame rate at 8.86 ± 0.14 Hz, and the tactile and IMU data were recorded at 32.0 ± 5.23 Hz.

Local geometric fidelity was quantified by the per-block prediction error ϵ_{block}, defined as the direct corner error of each 5-frame inter-block transformation ΔT against its GT within the sliding-window mechanism. This metric reflects short-range tracking precision. Global trajectory consistency was quantified by the accumulated tracking error ϵ_{acc} , computed across all block-representative frames to capture how local errors propagate throughout the scan duration. Long-term stability was assessed through the final drift ϵ_{drift}, defined as the Euclidean distance between the corners of the terminal frame and their GT positions.

TABLE I: Ablation study on sensor modalities. Rel. is error relative to the mean trajectory length (761.6 mm).

Metric	Config	Mean (mm)	Std (mm)	Rel. (%)
ϵ_{block}	Image	7.50	7.97	0.98
	+ Tactile	5.36	5.26	0.70
	+ Tactile & IMU	6.44	6.54	0.85
ϵ_{acc}	Image	101.30	86.62	13.30
	+ Tactile	50.17	29.73	6.59
	+ Tactile & IMU	33.91	19.31	4.45
ϵ_{drift}	Image	349.61	255.21	45.91
	+ Tactile	122.10	81.45	16.04
	+ Tactile & IMU	91.11	34.83	11.96

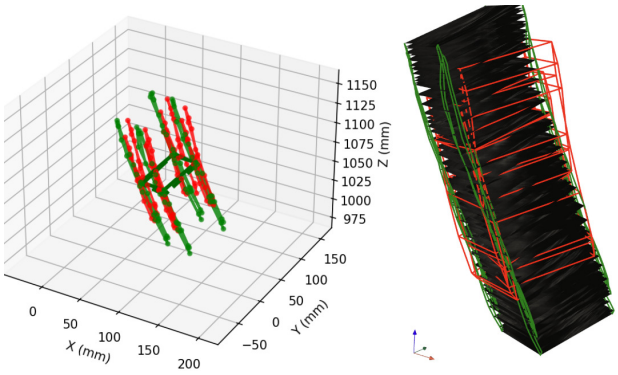


Fig. 3: Pose estimation overlay and reconstructed 3D US volume (trajectory length 893 mm). Red: predicted trajectory, Green: GT trajectory.

Table I reports the results for our full model and two ablated variants. The full model achieves a mean per-block error of 6.44 mm, accumulated error of 33.91 mm, and final drift of 91.11 mm, confirming reliable tracking under the random-walk accumulation. In contrast, the image-only baseline suffers severe drift of 349.61 mm, while tactile sensing alone reduces drift to 122.10 mm. Fusing IMU and tactile further lowers drift to 91.11 mm and yields the lowest accumulated error of 33.91 mm, demonstrating that multi-modal sensor fusion substantially improves long-range trajectory consistency. Fig. 3 illustrates a representative sequence where predicted and GT poses stay closely aligned throughout the bulk of the scan, with the largest deviations concentrated near the trajectory endpoints where accumulated drift is greatest.

CONCLUSIONS AND DISCUSSION

A multi-sensor fusion trackerless 3D freehand US reconstruction framework for orthopedic surgery has been presented, pairing the probe with low-cost inertial and tactile sensors to remove external tracking hardware. While the current accuracy does not yet match optical or robotic tracking, the system offers a simple, low-cost, and workflow-friendly route toward intraoperative 3D US navigation.

REFERENCES

[1] T. Liang, L. Yung, and W. Yu, "On feature motion decorrelation in ultrasound speckle tracking," *IEEE Transactions on Medical Imaging*, vol. 32, no. 2, pp. 435–448, 2013.

[2] Q. Li *et al.*, "Trackerless freehand ultrasound with sequence modelling and auxiliary transformation over past and future frames," in *2023 IEEE 20th ISBI*, 2023, pp. 1–5.

[3] B. Koonce, "Efficientnet," in *Convolutional neural networks with swift for Tensorflow: image recognition and dataset categorization*. Springer, 2021, pp. 109–123.

[4] Y. Zhou, C. Barnes, J. Lu, J. Yang, and H. Li, "On the continuity of rotation representations in neural networks," in *Proceedings of the IEEE/CVF CVPR*, June 2019.

Electrical Bio-Impedance Sensing on 2-Layered Tissue Model

Marianne V. Petersen, Thiusius R. Savarimuthu, and Zhuoqi Cheng

Maersk Mc-Kinney Moller Institute, SDU Robotics, University of Southern Denmark, Odense, Denmark

INTRODUCTION

Upper aerodigestive tract (UADT) cancers represent a substantial global health burden. Early detection is critical for improving patient outcomes; however, current diagnostic pathways remain limited in sensitivity at early disease stages. Standard clinical assessment relies on visual inspection using rhinolaryngoscopy, to examine the mucosal surfaces of the UADT. Suspicious lesions are subsequently evaluated through biopsy for histopathological confirmation. This approach is inherently subjective and depends heavily on the clinician's experience. Moreover, early-stage lesions are often subtle and can be obscured by inflammation, bleeding, or surface secretions, making reliable detection challenging [1].

Electrical Bio-Impedance (EBI) has emerged as a promising modality for tissue characterization and cancer detection, as it enables differentiation of biological tissues based on their electrical properties [2]. Building on this principle, a novel version of the SmartProbe has been developed with optimized electrode placement (Fig. 1) designed for characterization of layered mediums. The probe has a diameter of 1.8 mm and features a 45° angled tip to facilitate tissue contact. It integrates four gold electrodes, each with a diameter of 0.2 mm, arranged circumferentially around the probe tip with a defined angular separation. The SmartProbe is designed to be deployed through the working channel of a laryngoscope, allowing direct contact with suspicious tissue for in situ measurement of electrical properties.

The optimized electrode configuration enables depth-sensitive electrical sensing, allowing interrogation of tissue properties at controlled penetration depths [3]. This capability is particularly relevant for the head and neck region, where tissues exhibit a layered structure and where squamous cell carcinoma (SCC) primarily originates in the superficial layers, originating in the basal layer of the epithelium, spreading toward the surface before breaking through the basement membrane into the lamina propria. Targeting the superficial layers is thereforecritical for early detection and characterization of malignant changes. In this study, we present a method for measuring and estimating tissue conductivity using the SmartProbe, with a focus on capturing layer-specific electrical properties and providing experimental validation of the proposed approach.

MATERIALS AND METHODS

The tripolar impedance sensing method is for this purpose, since it allows for more unique measurement

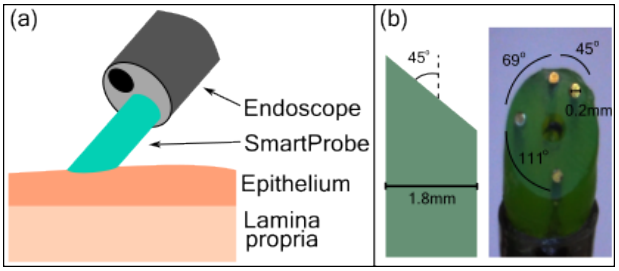


Fig. 1: (a) The SmartProbe is designed to pass through the endoscope working channel, and to contact UADT tissue surface for inspection; (b) The dimension of SmartProbe and the embedded position-optimized electrodes.

configurations with a limited number of electrodes [3]. One electrode, A, is used for current injection, while electrodes B and C are used for voltage measurements. When measuring a layered material, the current is partially reflected at the interface between the layers. The proportion of reflected current is determined by the conductivity contrast between the two layers, called the reflection coefficient ($r = \frac{\sigma_{ep} - \sigma_{lp}}{\sigma_{ep} + \sigma_{lp}}$), where σ_{ep} and σ_{lp} represent the conductivity of the epithelium layer and the lamina propria layer respectively. In this case, the measured conductivity, referred to as the apparent conductivity (σ_a), contains information about both layers. Theoretically, the apparent conductivity of a layered material is given:

$$\frac{1}{\sigma_a} = \frac{1}{\sigma_{ep}} + \frac{2}{\sigma_{ep}} \frac{d_{AC} \cdot d_{AD}}{|d_{AC} - d_{AD}|} \sum_{n=1}^{\infty} \left(\frac{r^n}{\sqrt{d_{AC}^2 + (2nh)^2}} - \frac{r^n}{\sqrt{d_{AD}^2 + (2nh)^2}} \right) \tag{1}$$

where h is the thickness, n is the number of reflections and d denotes the inter-electrode distances. Due to size constraints, only four electrodes are integrated into the SmartProbe. Their spatial arrangement is optimized such that the inter-electrode distances are non-uniform. By reusing each electrode for either current injection or voltage measurement, a total of 12 distinct tripolar sensing configurations can be realized. The variation in inter-electrode distances enables sensitivity to different penetration depths, thereby facilitating the extraction of depth-dependent electrical information and improving the reconstruction of layered tissue structures.

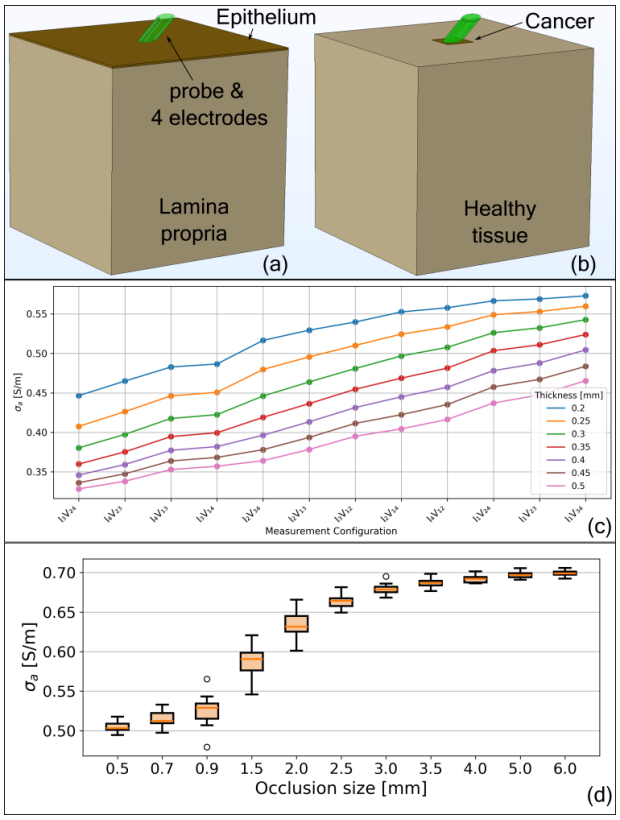


Fig. 2: Two simulation setups(a) 2-layered UADT tissue model, and (b) SCC cancer growing on the surface of healthy tissue; (c-d) the apparent conductivity results of setup (a) and (b), respectively. Variation in (d) represents the 12 tripoalar measurements

RESULTS

To evaluate the proposed probe, finite element simulations were conducted in *COMSOL Multiphysics 6.2*. Four surface electrodes, representing the SmartProbe configuration, were positioned on a 20 × 20 × 20 mm material. Two models were considered (Fig. 2): (a) two-layer UADT tissue and (b) SCC growing on healthy tissue.For model (a), the conductivities of the lamina propria and epithelium were set to 0.6 S/m and 0.3 S/m respectively.The epithelial thickness was varied between 0.2–0.5 mm to reflect anatomical variability. For model (b), the conductivity of the healthy tissue was set to 0.5 S/m, while the cancer was at 0.7 S/m. Different sizes of a cubic occlusion were added to the model aiming to assess its sensitivity of small lesions, conducting all 12 tripolar measurements for each. These simulations assessed measurement accuracy, depth sensitivity, and lesion detectability under idealized conditions.

For setup (a), the resulting apparent conductivity (σ_a) for each measurement configuration is shown in Fig. 2(c), sorted from most superficial to least superficial. I_AV_{BC} represents current injection through electrode A, and voltage measured between Electrode B and C. Different curves represent the results of different top layer thickness. These results demonstrate the probes ability to capture information from both superficial and deeper layers across a physiologically relevant range of layer thicknesses.

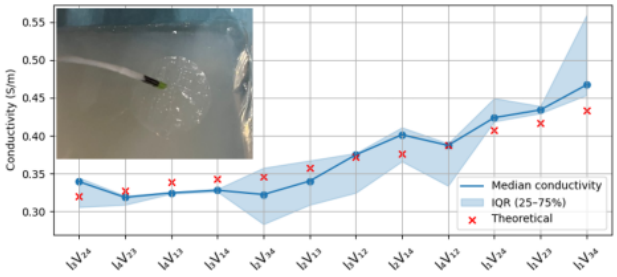


Fig. 3: The theoretical and measured values of 2-layered agar phantom: top layer 0.3 S/m (thickness 0.5 mm), and bottom layer 0.6 S/m.

The resulting apparent conductivity for setup (b) is shown in Fig. 2(d). When the lesion size is very small, such as ≤0.9 mm, the measurements reflect mainly the base material; while the values reflect the cancer when it is ≥3.5 mm. A transition region can also be seen from the results.

A more realistic setup is used for characterization. A phantom consisting of two distinct agar layers is fabricated to mimic the epithelial and lamina propria layers of the UADT. The top layer was produced by placing agar in a hand microtome and slicing a thin section. The target thickness of this layer was 0.5 mm, although the actual thickness may be slightly different and hard to be measured accurately due to the slicing tolerance. The electrical conductivities of the top and bottom agar layers are expected to be 0.3 S/m and 0.6 S/m, respectively.

The measurements on the agar phantom surface are repeated multiple times, and the results are presented in Fig. 3, alongside their theoretical values shown by red stars. The Interquartile Range (IQR) is also shown for each of the 12 tripolar configurations, indicating the spread in measured conductivity. The measured conductivity agrees with the theoretical values with an accuracy of 95.6%.

CONCLUSIONS AND DISCUSSION

These results demonstrate the probes ability to capture information from both superficial and deeper layers across a physiologically relevant range of layer thicknesses. In case of a thinner top layer, configurations with higher superficial sensitivity exhibit greater variation, whereas increasing thickness leads to greater variation in configurations with deeper sensitivity.

REFERENCES

[1] C. Sampieri, M. A. Azam, A. Ioppi, C. Baldini, S. Moccia, D. Kim, A. Tirrito, A. Paderno, C. Piazza, L. S. Mattos *et al.*, "Real-time laryngeal cancer boundaries delineation on white light and narrow-band imaging laryngoscopy with deep learning," *The Laryngoscope*, vol. 134, no. 6, pp. 2826–2834, 2024.

[2] A. L. C. Carobbio, Z. Cheng, T. Gianiorio, F. Missale, S. Africano, A. Ascoli, M. Fragale, M. Filauro, F. Marchi, L. Guastini *et al.*, "Electric bioimpedance sensing for the detection of head and neck squamous cell carcinoma," *Diagnostics*, vol. 13, no. 14, p. 2453, 2023.

[3] M. V. Petersen, T. R. Savarimuthu, and Z. Cheng, "Development of a bioimpedance probe for enhanced tissue identification in the upper aerodigestive tract," in *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2025, pp. 1–6.

Feasibility of Machine Learning-Based Tissue Discrimination Using Electrical Bioimpedance Sensing

Yao Zhang*[1], Anas El Hila*[1], Teshome Getu Mhret[1], Mouloud Ourak[1], and Emmanuel Vander Poorten[1]

[1]*Department of Mechanical Engineering, KU Leuven, Belgium*

INTRODUCTION

Tissue discrimination during minimally invasive interventions is critical for improving procedural safety and therapeutic precision. In many catheter- and needle-based procedures, direct visualization of the instrument tip is limited or unavailable due to restricted access and complex anatomical environments. Consequently, accurately identifying the surrounding tissue in real time remains a major clinical challenge. For example, during femoral artery access (FAA), distinguishing tissue layers while advancing the needle is essential to avoid complications such as posterior wall perforation [1]. Although ultrasound enables real-time visualization, reliable tissue discrimination still requires substantial operator training, while image quality and spatial resolution may remain insufficient in challenging anatomical conditions [2]. Similarly, in endovascular interventions for thrombotic occlusions, identifying thrombus composition, such as red blood cell (RBC)-rich or fibrin/platelet-rich thrombi, is important for selecting appropriate thrombectomy or thrombolytic strategies and improving treatment outcomes [3]. However, current thrombus characterization methods primarily rely on laboratory analysis or post-procedural histopathology, which are time-consuming and limit real-time clinical decision-making [4].

Electrical bioimpedance (EBI) has emerged as a promising sensing modality for minimally invasive procedures due to its low cost, compact integration capability, and sensitivity to tissue electrical properties. By analyzing the electrical response around the sensing electrode, EBI can capture variations in tissue conductivity and permittivity associated with different biological tissues [5]. However, conventional approaches often rely on manually selected impedance features or physics-based analytical models, which may have limited generalizability across complex tissue environments [6]. To address this, the feasibility of combining EBI with machine learning (ML)-based methods for tissue discrimination was investigated in this work, aiming to enable more adaptive and robust tissue classification.

MATERIALS AND METHODS

Three machine learning (ML)-based classifiers, namely logistic regression (LR), support vector machine (SVM), and partial least squares discriminant analysis (PLS-DA) [7], were investigated in this work to evaluate the feasibility of electrical bioimpedance (EBI)-based tissue discrimination and compare classification performance. To assess tissue discrimination capability, two different *ex vivo* biological tissues were selected: pork muscle and pork fat, representing two distinct tissue classes with different electrical properties. As shown in Fig. 1, ten samples of pork muscle and ten samples of pork fat were prepared. To ensure measurement consistency, all tissue samples were adjusted to a similar mass using a digital scale before the experiments. Each tissue sample was placed inside a measurement tube, where a SMARTTOUCH® catheter was positioned above the tissue using a motorized setup. The distal ring electrodes at the catheter tip were configured in a tetrapolar arrangement for EBI measurement [8]. During the experiment, the catheter was advanced towards the tissue with constant speed via a DC motor (315685, Maxon Group, Switzerland) until the catheter tip established contact with the tissue surface.

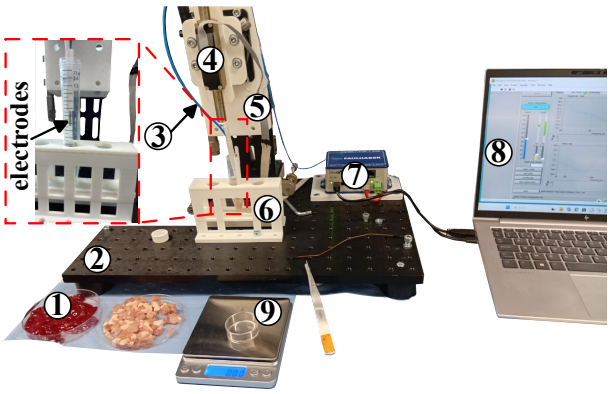


Fig. 1: ①–Tissue samples; ②–Breadboard; ③–Catheter; ④–DC motor; ⑤–Multi-DOF stage; ⑥–Tube holder; ⑦–Motor controller; ⑧–EBI sensor GUI; ⑨–Digital scale

During each measurement, alternating current (AC) excitation signals at 15 different frequencies (1 kHz - 349 kHz) were applied for impedance acquisition during tissue contact. For each sample, both impedance magnitude and phase were recorded across all excitation frequencies with a sampling frequency of 1 kHz, resulting in 15 magnitude features and 15 phase features. Consequently, each measurement produced a 30-dimensional feature vector that served as the input for the ML-based classifiers. The acquired impedance signals were preprocessed using a moving average filter with a window size of 50 samples to reduce measurement noise. Specifically, for every consecutive 50

*These authors contributed equally to this work.

This work was supported in part by the European Union through the AIRCARE innovation action project (No. 101137426), and in part by the C2 project RIVUS (C2E/25/024), and in part by the PROMPT project from the European Innovation Council (GA101291374).

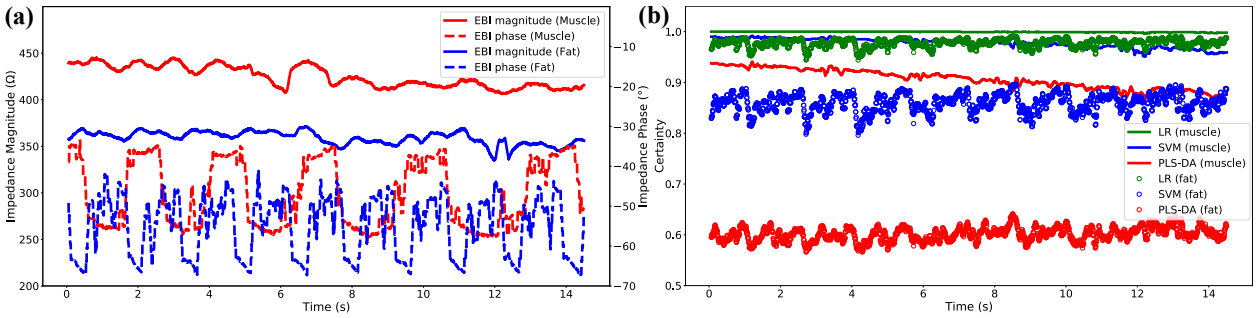


Fig. 2: Tissue discrimination results based on EBI measurements. (a) Representative impedance measurement data at a selected excitation frequency (1 KHz). (b) Tissue classification certainty obtained using LR, SVM, and PLS-DA algorithms for pork muscle and fat tissue discrimination.

TABLE I: Quantitative tissue discrimination results for pork muscle and fat tissues

Metric / Algorithms / Tissue	LR		SVM		PLA-DA	
	Mean	SD	Mean	SD	Mean	SD
Muscle	0.99	0.0008	0.98	0.0098	0.90	0.0202
Fat	0.98	0.0080	0.86	0.0202	0.60	0.0189

data points, the mean values of the impedance magnitude and phase at each excitation frequency were calculated and used to form the input feature vector for the ML-based classifiers. In total, 20 groups of data (10 muscle and 10 fat samples) were used as the training dataset. To evaluate classifier performance, stratified 4-fold cross-validation was performed to maintain balanced tissue classes across folds while ensuring a fair comparison among the three classifiers. After model training, one additional pork muscle sample and one additional pork fat sample, not included in the training dataset, were measured and used as independent test samples for tissue discrimination validation.

RESULTS

Fig. 2 presents the tissue discrimination results obtained using the three investigated machine learning classifiers. Fig. 2(a) shows representative test data of the impedance magnitude and phase responses of pork muscle and fat tissues at a selected excitation frequency (1 KHz) over time. Fig. 2(b) presents the classification certainty obtained using the three investigated machine learning classifiers. Overall, LR achieved the highest and most stable classification certainty throughout the measurement duration, followed by SVM, whereas PLS-DA showed comparatively lower certainty values. In addition, larger variations were observed for fat tissue, particularly for SVM and PLS-DA, indicating increased classification difficulty for this tissue type.

The quantitative evaluation results are summarized in Table I. The mean values correspond to the average classification certainty over time, while the standard deviation (SD) represents the variability across repeated classifications obtained from individual 50-sample averaging windows. For muscle discrimination, LR achieved the highest mean certainty of 0.99 with a standard deviation (SD) of 0.0008, followed by SVM with a mean certainty of 0.98 and SD of 0.0098, while PLS-DA achieved a mean certainty of 0.90 with SD of 0.0202. Similar trends were observed for fat tissue classification, where LR achieved the best performance and PLS-DA showed the lowest performance.

CONCLUSIONS AND DISCUSSION

In this work, the feasibility of combining EBI sensing with ML-based methods for tissue discrimination was investigated using *ex vivo* pork muscle and fat tissues. Three classifiers, namely LR, SVM, and PLS-DA, were evaluated and compared. The experimental results demonstrated that EBI measurements acquired from the catheter tip enable feasible and reliable tissue classification.

The obtained results showed that muscle tissue generally achieved higher classification certainty compared with fat tissue. One possible explanation is that muscle tissue has higher water and ionic content, resulting in stronger conductivity and more stable impedance characteristics across frequencies. Furthermore, the superior performance of LR may be related to the relatively small dataset used in this preliminary study. Since only 20 training samples were available, simpler classifiers may generalize better and be less prone to overfitting compared with more complex classifiers. In contrast, SVM and PLS-DA may require larger datasets to fully exploit their classification capabilities and achieve more robust decision boundaries. Future work will focus on expanding the dataset size, investigating additional tissue types, and exploring more advanced ML and deep learning approaches for real-time tissue characterization during minimally invasive interventions.

REFERENCES

[1] E. M. Campion and C. J. Fox, "Prehospital hemorrhage control and reboa," *Current Trauma Reports*, vol. 5, pp. 129–136, 2019.

[2] M. Afshar *et al.*, "Autonomous ultrasound scanning to localize needle tip in breast brachytherapy," in *2020 International Symposium on Medical Robotics (ISMR)*. IEEE, 2020, pp. 202–208.

[3] W. Brinjikji *et al.*, "Correlation of imaging and histopathology of thrombi in acute ischemic stroke with etiology and outcome: a systematic review," *Journal of neurointerventional surgery*, vol. 9, no. 6, pp. 529–534, 2017.

[4] S. Staessens *et al.*, "Structural analysis of ischemic stroke thrombi: histological indications for therapy resistance," *haematologica*, vol. 105, no. 2, p. 498, 2019.

[5] H. Kalvøy *et al.*, "Impedance-based tissue discrimination for needle guidance," *Physiological measurement*, vol. 30, no. 2, p. 129, 2009.

[6] Y. Zhang *et al.*, "Electrical bioimpedance sensing-based proximity distance estimation for needle interventions: Two-layer model and validation," *IEEE Trans. on Instrumentation and Measurement*, 2026.

[7] A. F. Alnuaimi and T. H. Albaldawi, "An overview of machine learning classification techniques," in *BIO Web of Conferences*, vol. 97. EDP Sciences, 2024, p. 00133.

[8] Y. Zhang *et al.*, "Electrical bioimpedance sensing-based micro-catheter alignment in cto-pci," *IEEE Sensors Journal*, vol. 24, no. 19, pp. 31 082–31 094, 2024.

Evaluating Kinematic Tracking Stability and Load-Dependent Error in the da Vinci Research Kit (dVRK)

Toilos Christos, Mazomenos Evangelos, and Mennillo Laurent

University College London, Department of Medical Physics and Biomedical Engineering

INTRODUCTION

Autonomous robotic surgery, such as automated tissue dissection, requires high-fidelity spatiotemporal tracking for patient safety [1]. Cable-driven surgical robots like the da Vinci Research Kit (dVRK) calculate end-effector positions using internal encoders and forward kinematics. However, tracking precision is limited because these models assume infinite joint stiffness and rigid linkages. Consequently, unmodeled structural compliance, cable hysteresis, and joint uncertainties create significant discrepancies between commanded and physical positions, especially under variable tissue resistance [2], [3]. To correct this, researchers rely on camera-robot (hand-eye) calibration to establish rigid transformations between the endoscope and robot base [4], often incorporating Remote Center of Motion (RCM) constraints [5]. While static calibration minimizes baseline errors in free space, its dynamic performance under continuous operation and load-bearing tissue resistance remains largely unquantified. To inform robust control system design, this study uses optical tracking to investigate the dVRK's spatiotemporal kinematic integrity. We evaluate two primary objectives: (1) determining if progressive temporal drift occurs over prolonged, repetitive trajectories, and (2) quantifying spatial accuracy degradation induced by external forces. This study evaluates the dynamic spatiotemporal stability of the joint tracking-and-manipulator pipeline across continuous multi-dimensional trajectories under simulated payloads. By examining baseline stationarity and force-induced divergence on a unified physical platform, we provide a characterization necessary to establish the operational boundaries of offline hand-eye calibration before deploying real-time correction.

MATERIALS AND METHODS

1) Experimental Setup: The experimental platform comprised a da Vinci Research Kit (dVRK) system (Fig. 1). A Patient Side Manipulator (PSM1) equipped with a Large Needle Driver (LND, model 400006) provided actuation, while the Endoscopic Camera Manipulator (ECM) captured visual data. Monocular tracking utilized the endoscope's left optical channel at 15 frames per second. To quantify spatial discrepancies between internal kinematics and physical reality and evaluate the accuracy degradation of the integrated tracking-and-manipulator system, a 10×10 mm ArUco marker (`DICT_4X4_50`) was attached to the LND's distal shaft. To simulate *in vivo* tissue resistance, a multi-directional force field was created using elastic rubber bands anchored between the instrument shaft and the surrounding workspace. Rather than a calibrated force model measured in Newtons, this setup provides a qualitative, step-wise load profile to observe tracking performance under resistance.

2) Data Collection: The PSM1 executed a repeatable 3-dimensional Figure-8 (Lissajous) trajectory using a stop-and-go profile, discretized into 16 spatial waypoints per loop. Two experimental paradigms were tested: (1) free-space motion (control) and (2) resisted motion (with elastic bands). Approximately 1,000 discrete waypoint measurements were recorded per condition. During resisted motion, High-Current Quad Linear Amplifier (QLA) thermal warnings required temporarily releasing elastic tension to prevent hardware throttling. Two bands were initially detached to reduce actuator load and sequentially re-engaged once thermal equilibrium was restored.

3) Algorithms and Error Metrics: The software pipeline computed the discrepancy between internal kinematic estimations and camera-observed physical ground-truth. First, the dVRK API queried the robot's forward kinematics for the jaw tip's 3D pose relative to the base frame. Because the camera tracks the ArUco marker, a static 4×4 rigid-body offset matrix mapped this tip estimate directly to the marker's center. This base-frame coordinate was then transformed into the camera's optical frame via a pre-computed Hand-Eye calibration matrix. Simultaneously, the marker's physical ground-truth position was extracted from the endoscope feed using OpenCV-based ArUco detection adapted from the `dvrk_camera_registration` repository [6].System accuracy was quantified at each waypoint using two metrics:

3D Euclidean Error (mm): The absolute 3-dimensional spatial distance between the kinematic prediction (P_{est}) and visual reality (P_{actual}).

2D Reprojection Error (pixels): The Euclidean pixel distance between the visually detected marker centroid and the 3D kinematic estimate (P_{est}) projected backwards onto the 2D image plane using the camera's intrinsic distortion matrix.

CONCLUSIONS AND DISCUSSION

Baseline Kinematic Tracking: Under free-space conditions, baseline average errors were 10.87 ± 6.19 mm (3D Euclidean) and 4.67 ± 1.18 pixels (2D reprojection) (Fig. 2, Bottom). High-frequency oscillatory error profiles maintained a constant mean across waypoints. Longitudinal analysis revealed no progressive drift, indicating stable robot-

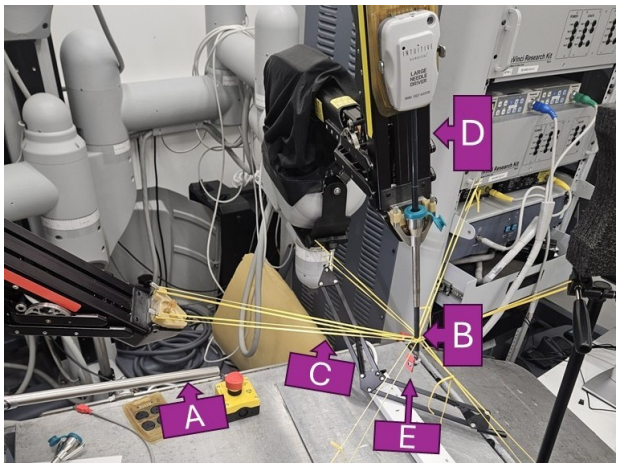


Fig. 1: Overview of the experimental dVRK robotic setup indicating key components. **(A)** Endoscopic Camera Manipulator (ECM), **(B)** Large Needle Drive (LND), **(C)** Distal elastic bands applying multi-directional force field, **(D)** Patient Side Manipulator (PSM1), **(E)** 10 mm ArUco marker attached to the end-tip of the LND.

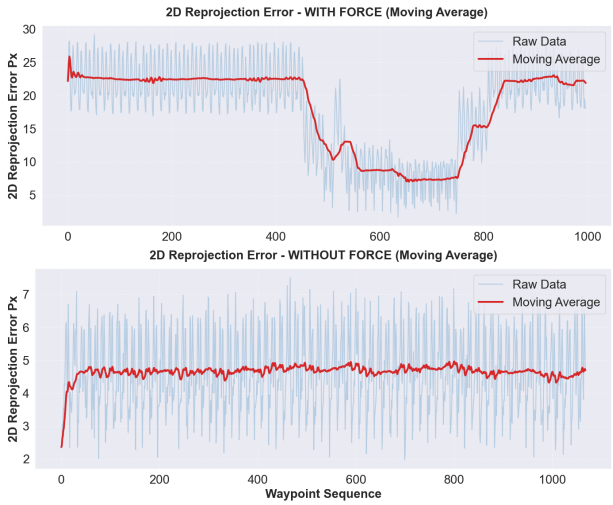


Fig. 2: Overall performance of the kinematic tracking. **(Top)** 2D Reprojection Error (in pixels) with force applied. **(Bottom)** 2D Reprojection Error (in pixels) without force applied.

to-camera transformation without cumulative encoder drift. An Augmented Dickey-Fuller (ADF) test confirmed strictly stationary baseline 3D error ($p < 0.0001$), establishing non-decaying variance over time. Crucially, this baseline average error of ~10.87 mm is not a direct reflection of the robot's physical repeatability; rather, it represents the cumulative error of the entire integrated system, which includes significant observational and calibration artifacts. These stem from monocular ArUco tracking depth uncertainty, hand-eye calibration tolerances, and physical marker offset estimation.

Effects of External Force on Kinematic Accuracy: Simulated tissue resistance significantly degraded tracking accuracy (Welch's t-test, $p < 0.0001$), increasing average 3D error to 17.83 ± 5.72 mm and 2D reprojection error to 18.07 ± 6.89 pixels (Fig. 2, Top). The integrated tracking-and-manipulator system proved highly sensitive to external payloads. While maximum loads elevated errors consistently (Fig. 2, Top), temporarily detaching two elastic bands (approx. waypoint 450) plummeted the 2D error from ~23 to 8 pixels. Sequentially re-applying bands (waypoints ~750 and ~800) caused distinct, step-wise error increases.

Discussion: These findings highlight two critical limitations of the integrated tracking-and-manipulator pipeline. First, broad free-space tracking variance (3D error SD = 6.19 mm) reveals spatially dependent hand-eye calibration and monocular tracking errors. Crucially, this monocular setup cannot cleanly isolate force-induced physical deflections from view-dependent tracking artifacts. Second, the system is highly vulnerable to unmodeled structural compliance (e.g., cable stretch, pulley backlash) ignored by standard rigid kinematic models. External tip forces thus cause physical deflections unrecorded by motor encoders. While our repeatable elastic loads yielded steady average errors without continuous drift, chaotic *in vivo* surgical forces will likely exacerbate these discrepancies, further necessitating online hand-eye calibration.

Conclusion: This study evaluated dVRK PSM tracking stability within the integrated system. While trajectories exhibited no progressive temporal drift, spatial accuracy remained highly susceptible to external forces. Qualitative simulated tissue resistance increased average 3D tracking error by ~64% (10.87 mm to 17.83 mm), with manual force adjustments causing immediate step-wise shifts. These results underscore the necessity for online hand-eye calibration via real-time visual tracking to achieve precision during autonomous manipulation. Future work will incorporate stereo tracking and calibrated load sensors to cleanly isolate manipulator compliance from observational artifacts.

REFERENCES

[1] D. A. Yip *et al.*, "Autonomous dissection in robotic cholecystectomy," *arXiv preprint arXiv:2503.00666*, 2025.

[2] Z. Cui, J. Cartucho, S. Giannarou, and F. Rodriguez y Baena, "Caveats on the first-generation da vinci research kit: Latent technical constraints and essential calibrations [survey]," pp. 113–128, 2025.

[3] B. Li, H. Lin, F. Zhong, and Y. Liu, "Real-time geometric joint uncertainty tracking for surgical automation on the dvrk system," in *2024 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2024, pp. 2144–2150.

[4] O. Özgüner, T. Shkurti, S. Huang, R. Hao, R. C. Jackson, W. S. Newman, and M. C. Çavuşoğlu, "Camera-robot calibration for the da vinci robotic surgery system," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 2154–2161, 2020.

[5] R. Wei, B. Li, H. Mo, F. Zhong, Y. Long, Q. Dou, Y.-H. Liu, and D. Sun, "A unified monocular camera-based and pattern-free hand-to-eye calibration algorithm for surgical robots with rcm constraints," *IEEE Robotics and Automation Letters*, vol. 6, p. 5124, 2022.

[6] Johns Hopkins University LCSR, "da vinci research kit camera registration," https://github.com/jhu-dvrk/dvrk_camera_registration, 2026.

Pneumatically Actuated MRI-Compatible Robot for Spinal Needle Procedures

Yuxuan Qi[1], Aminul I Ahmed[2], Christos Bergeles[3] and Zion Tse[1]

[1] *School of Engineering and Materials Science, Queen Mary University of London,*

[2] *Department of Neurosurgery, King's College Hospital NHS Foundation Trust*

[3] *School of Biomedical Engineering & Imaging Sciences, King's College London*

z.tse@qmul.ac.uk

INTRODUCTION

Spinal needle procedures, including spinal nerve root injections and lumbar punctures, are widely used for diagnosis and treatment. [1] In these procedures, accurate needle placement is required to reach the target region while avoiding sensitive anatomical structures. Magnetic resonance imaging (MRI) provides high spatial resolution and excellent soft-tissue contrast, which is valuable for spinal diagnosis and image-guided intervention. However, robotic assistance under MRI remains challenging because of the strong magnetic field, radiofrequency pulses, and gradient fields. Conventional electromagnetic actuators are generally unsuitable near the imaging region, while manual needle alignment inside the confined MRI environment still relies heavily on clinical experience. Li et al. developed a piezoelectrically actuated, body-mounted robotic system for MRI-guided lower-back injections, demonstrating the feasibility of robotic needle alignment in this application [2]. Squires et al. presented an MR-safe pneumatically actuated spinal robot, in which the electromagnetic control components were located remotely from the patient-mounted mechanism [3]. However, structural compactness and angular workspace remain important design considerations for lower-back spinal injections.

The present study proposes a compact 2-DOF pneumatic robot that provides pitch and yaw adjustment of the needle guide before insertion. The system incorporates a rotary disc motor for pitch adjustment and a piston-based stepper motor for yaw adjustment, enabling active motion control without electromagnetic actuation near the imaging region.

MATERIALS AND METHODS

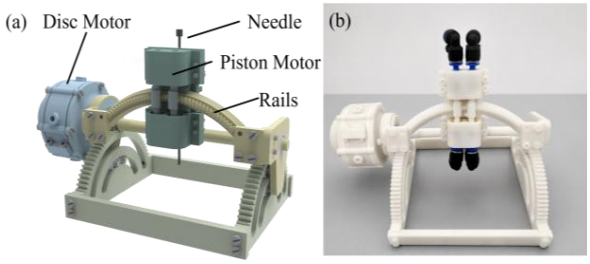


Figure. 1. Overall design and fabricated prototype of the 2-DOF pneumatically actuated spinal injection robot.

As shown in Figure. 1, we built a pneumatically driven 2-DOF robot. For the base section, a rotary disc motor is fitted on one side of two parallel semicircular rails, which provides the pitch movement. For the upper part, one arc-shaped rail moves together with the disc motor, and a piston motor fitted on this rail can realize stepper yaw movement. The insertion needle can pass through a fixed hole at the back of the piston motor.

The robot was designed for MRI compatibility using entirely non-metallic components. The frame structure was fabricated using photopolymer resin with a Formlabs Form 4 printer, and all connecting components were made of plastic materials. The fabricated prototype was measured to have an overall size of 105 mm × 115 mm × 104.3 mm. The available angular motion ranges were approximately ±38° in pitch and ±45° in yaw. These measured geometric parameters were used to build a MATLAB-based workspace model for estimating the reachable insertion region.

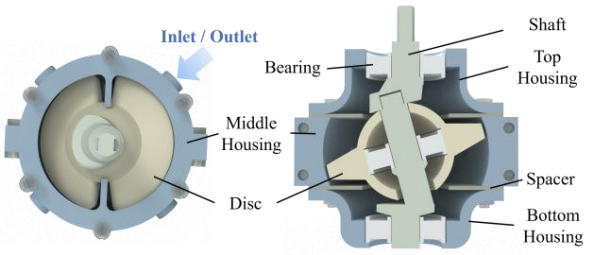


Figure. 2. CAD structure of the rotary disc motor.

The disc motor is adapted from the nutating disc engine.[4] Additional internal partitions were introduced to divide the cavity into multiple chambers and constrain disc motion, enabling sequential pressure control through four inlet/outlet ports operated by external valves. In preliminary testing, the motor reached 900 rpm at 400 kPa.

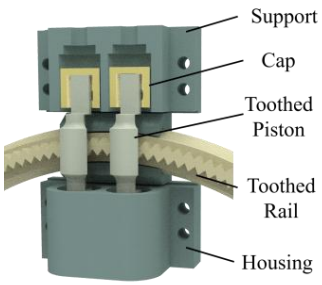


Figure. 3. Structure of the piston-based stepper motor.

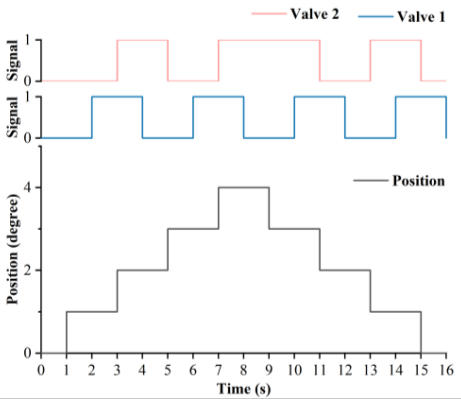


Figure. 4. Control sequence and stepwise angular response of the piston motor.

As shown in Figure. 3, the piston motor includes two opposed toothed pistons. A cylindrical cap is mounted at the end of each toothed piston, allowing it to slide inside the chamber. The piston motor contains four chambers. In the stationary state, two toothed pistons are always engaged with the rail structure, providing a self-locking function in place. Two alternating piston movements generate one 1° output step. Sequential pressurization of the four chambers enables bidirectional motion. Figure 4 shows the valve signals and angular response, with each valve phase held for 1 s.

RESULTS

Based on the measurements of the robot, MATLAB was used to simulate the workspace of the spinal injection robot. The model used the measured angular limits of ±38° in pitch and ±45° in yaw, with a maximum insertion depth of 90 mm.

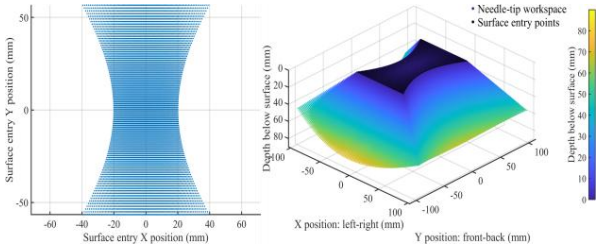


Figure. 5. Simulated insertion workspace of the spinal robot with a maximum insertion depth of 90 mm.

As shown in Figure. 5, although the two DOFs can be actuated separately, the fixed needle guide still causes coupling between the two directions. Therefore, the reachable insertion range along the y-axis is not linear, but forms a concave boundary. This phenomenon is not observed along the x-axis because the motion is constrained by the base frame and does not reach the theoretical boundary. In the physical measurement, due to obstruction from the two motor structures, the overall reachable motion range was approximately 67 mm × 65 mm.

To preliminarily assess positioning performance, accumulated angular displacement was determined from position marks on a 1-mm reference grid and compared with the designed position. As shown in Figure 6, larger deviations occurred near the middle of the rail, while the maximum absolute deviation was approximately 1.4°.

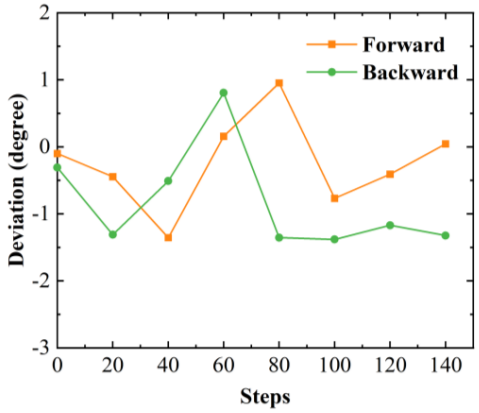


Figure. 6. Angular deviation of the piston motor during stepwise motion.

CONCLUSION AND DISCUSSION

This study presents a pneumatically actuated MRI-compatible robot for spinal needle procedures. The system uses two distinct pneumatic motor designs to provide pitch and yaw adjustment of the needle guide. The compact prototype demonstrated a useful angular motion range, stepwise yaw movement, self-locking capability, and a simulated reachable workspace for lower-back spinal intervention.

The current system is still at an early stage and requires further optimization. The rotary disc motor shows potential for high-speed pneumatic rotation, but its controllability, torque output, sealing performance, and positioning accuracy need further investigation. Future work will focus on improving motor performance, expanding the workspace, adding a reliable fixation method, validating needle alignment accuracy, and quantifying the robot's effects on MRI SNR and image artifacts.

REFERENCES

[1] L. Zhao, L. Manchikanti, A. D. Kaye, and A. Abd-Elsayed, 'Treatment of Discogenic Low Back Pain: Current Treatment Strategies and Future Options—a Literature Review', Curr Pain Headache Rep, vol. 23, no. 11, p. 86, Nov. 2019, doi: 10.1007/s11916-019-0821-x.

[2] G. Li et al., 'A Fully Actuated Body-Mounted Robotic Assistant for MRI-Guided Low Back Pain Injection', in 2020 IEEE International Conference on Robotics and Automation (ICRA), May 2020, pp. 5495–5501. doi: 10.1109/ICRA40945.2020.9197534.

[3] A. Squires, S. Hovet, R. Li, J. Oshinski, and Z. T. Ho Tse, 'A body-mounted device for MRI-guided spinal therapy', The International Journal of Medical Robotics and Computer Assisted Surgery, vol. 17, no. 2, p. e2235, 2021, doi: 10.1002/rcs.2235.

[4] T. Korakianitis, L. Meyer, M. Boruta, and H. E. McCormick, 'Alternative Multinutating Disk Engine Configurations for Diverse Applications', J. Eng. Gas Turbines Power, vol. 126, no. 3, pp. 482–488, Aug. 2004, doi: 10.1115/1.1688770.

Vibroacoustic-Based Bone Density Differentiation During Robotic Drilling

M. Van den Bogaert[1,*], M. Timmermans[1,4], K. Van Assche[2], A. Davoodi[2], W. Desmet[3,4], E. Vander Poorten[2], K. Denis[1]

[1]*Smart Instrumentation Research Team, Dpt. of Mechanical Engineering, KU Leuven, Belgium,* [2]*Robot-Assisted Surgery Group, Dpt. of Mechanical Engineering, KU Leuven, Belgium,* [3]*LMSD, Dpt. of Mechanical Engineering, KU Leuven, Belgium,* [4]*Flanders Make@KU Leuven, Belgium*

** Correspondence: marthe.vandenbogaert@kuleuven.be*

INTRODUCTION

Posterior pedicle drilling is routinely performed in spine surgery to place screws for stabilization in spinal trauma, deformity, and degenerative disc disease [1]. However, the procedure remains challenging due to the complex vertebral anatomy, proximity to critical neurovascular structures, and limited intraoperative feedback on instrument position, increasing the risk of cortical breach and tissue injury [2]. Variability in bone quality and limited intraoperative assessment methods further complicate prediction of screw fixation. Current guidance relies on radiation-intensive, preoperative imaging and navigation systems with limited real-time feedback [3]. Multimodal intraoperative sensing could improve robustness, reduce radiation, and enhance safety and efficacy [2]. Inspired by the surgeons' use of auditory and tactile cues during manual drilling, previous studies have explored audio and vibration sensing for state recognition during bone drilling [4] and milling [5], and breakthrough detection [6,7], including integration with robotic orthopedic systems. While these methods differentiate cortical and trabecular bone, local bone quality assessment remains unexplored. Therefore, this study investigates the potential of vibroacoustic signals from drill-bone interaction for bone quality assessment. Signals acquired during robotic drilling of artificial bone with varying densities are analyzed to discriminate bone density, and assess drilling parameter effects. A density resolution of 5 pcf is targeted to differentiate within and between healthy and osteoporotic bone. The results of this work support the development of real-time bone quality assessment and tissue differentiation methods for integration into multimodal guidance frameworks for minimally invasive and robot-assisted spine surgery.

MATERIALS AND METHODS

Artificial bone substitutes made of solid polyurethane (PU) foam (Sawbones, Vashon Island, WA, USA) with densities of 10 and 15 pcf were drilled using a KUKA Med7 robotic arm equipped with a custom drilling system. The PU blocks (30×30×40 mm) were fixed in a swivel bench vice, and holes 30 mm deep were drilled at controlled spindle speeds of 1500 or 2500 rpm, and feed rates of 0.5 or 2 mm/s. Five drilling trials were performed per parameter combination, leading to a total of 40 drillings. Lubrication (WD-40) was applied to the drill bit before each drilling. During drilling, vibroacoustic signals were acquired using different sensors: a PCB free-field condenser microphone (model 378B02, PCB Piezotronics, Depew, NY, USA) was positioned at approximately 10 cm from the drilling site, a triaxial accelerometer (model J356A45, PCB Piezotronics, Depew, NY, USA) was mounted on the drilling system with the z-axis aligned according to the drilling direction, and a lightweight uniaxial accelerometer (model 352A24, PCB Piezotronics, Depew, NY, USA) was attached to the PU samples at a distance of 1 cm from the drilling site using beeswax. All PCB sensors were connected via BNC cables to a custom signal conditioner. A Polytec IVS-500 laser vibrometer and corresponding junction box were installed to provide non-contact vibration measurements of the drill bit in radial direction. Vibroacoustic signals from all sensors were recorded at a sampling frequency of 40 kHz and automatically synchronized using a PicoScope 4824A oscilloscope (Pico Technology, Cambridgeshire, UK) and the corresponding PicoScope6 software. Figure 1 shows the experimental set-up.

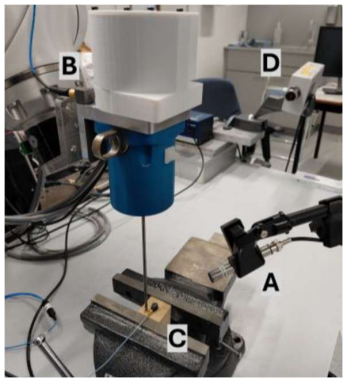


Figure. 1. Robotic drilling set-up and sensor configuration: (A) free-field microphone, (B) triaxial accelerometer, (C) uniaxial accelerometer, (D) laser vibrometer.

Vibroacoustic data from all channels were processed in MATLAB R2024b. Drill-sample contact onset was automatically detected from the uniaxial accelerometer signal. A 1 to 2 cm drilling segment after contact was extracted and normalized to its maximum amplitude. Power spectral densities (PSD) were computed using Welch's method with 4096-sample Hanning windows, 80% overlap, and a 4096-point Fast Fourier Transform. The PSDs were compared across conditions (free-running, 10 pcf, 15 pcf) to identify discriminative

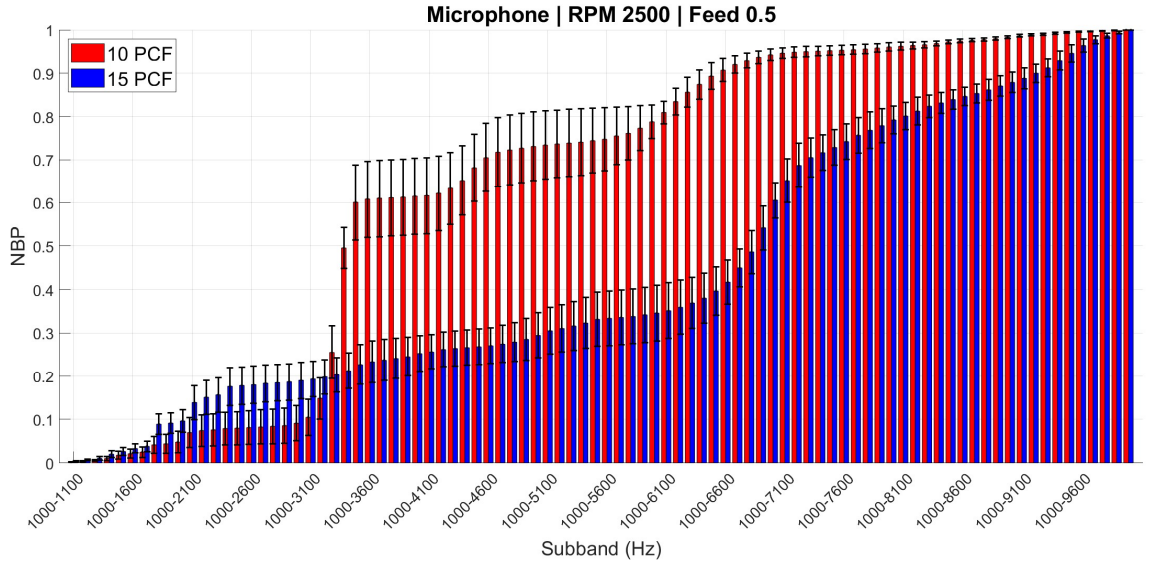


Figure. 2. Normalized Band Power (NBP) per frequency subband for 10 (red) vs. 15 (blue) pcf of the microphone signal for 2500 rpm drill speed and 0.5 mm/s feed rate.

frequency regions. Analysis focused on 1-10 kHz. Normalized Band Power (NBP) features were extracted from the PSD [8]. NBP is defined as the ratio of the power within a given frequency subband to the total spectral power:

$$NBP = \frac{\sum_{f_1}^{f_2} PSD(f)}{\sum_{f_1}^{f_3} PSD(f)} \cdot 100\ [\%]$$

where f_1=1 kHz, f_3=10 kHz, and f_2 is incremented from f_1 to f_3 in 100 Hz steps. For each sensor and drilling condition, bar charts compared NBP values for 10 and 15 pcf across all subbands. These analyses evaluated density discrimination, identified informative frequency bands, assessed sensor performance, and investigated drilling parameter effects. Global bar charts additionally compared NBP values for 10 and 15 pcf regardless of drilling parameters.

RESULTS

Significant differences between 10 and 15 pcf were observed for the microphone and uniaxial accelerometer in mid-high frequency subbands ($f_2 \approx$ 3.5 - 9.9 kHz, Mann-Withney U test, p<0.01), with higher NBP for lower-density foam. Figure 2 shows an example bar chart for microphone data. Global charts confirm that NBP-based discrimination remains effective across varying RPM and feed rates. For the laser vibrometer, NBP was higher for 15 pcf at low frequencies ($f_2 \approx$ 1.1 - 3 kHz, Mann-Withney U test, p<0.01) and higher for 10 pcf at mid-high frequencies ($f_2 \approx$ 4.5 - 9.9 kHz, Mann-Withney U test, p<0.01), enabling differentiation. The triaxial accelerometer showed no clear trend. Improved subband selection or sensor placement may improve performance. Overall, discrimination improved at high RPM and low feed rates, although further measurements are required. Note that based on visual PSD inspection, two measurement outliers were excluded from the analysis.

DISCUSSION AND CONCLUSION

The results show that vibroacoustic signals can distinguish material densities. NBP values in specific frequency bands correlated with density for the microphone, laser and uniaxial accelerometer, but not for the triaxial accelerometer. Density differences of 5 pcf were detectable, enabling discrimination between cortical and trabecular bone, and within trabecular bone. These findings support vibroacoustic sensing for real-time bone quality assessment. A limitation is that testing was performed in vitro using PU blocks rather than biological tissue. However, the evaluated sensor configuration was previously used successfully for breach detection in human cadavers [7]. Another limitation is the small sample size. Future work will expand the dataset, identify complementary features for a robust vibroacoustic signature across drilling conditions, and evaluate optimal sensing configurations. This will support development of an algorithm for automatic real-time bone quality assessment.

REFERENCES

[1] Perna F., et al. (2016). DOI: 10.1007/s12306-016-0438-8; [2] Timmermans M., et al. (2023). DOI : 10.3390/s23198094; [3] Yamout T., et al. (2023). DOI: 10.1016/j.ocl.2022.11.008; [4] Li M., et al. (2021). DOI : 10.1109/RCAR52367.2021.9517529; [5] Dai Y., et al. (2018). DOI: 10.1109/TMECH.2018.2804950; [6] Torun Y., et al. (2020). DOI : 10.1109/RCAR52367.2021.9517529; [7] Massalimova A., et al. (2023). DOI : 10.1016/j.artmed.2023.102641; [8] Goossens Q., et al. (2020). DOI: 10.48465/fa.2020.1132

ACKNOWLEDGEMENTS

FAROS has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016985.

Design of Adjustable Constant-Force Surgical Forceps for Tissue Handling

Haopeng Lu[1,2] , Zhuoqi Cheng[1], Thiusius R. Savarimuthu[1], and Di Wu[1]

1. The Maersk McKinney Moller Institute, University of Southern Denmark, Denmark

2. School of Engineering and Design, Technical University of Munich, Germany

INTRODUCTION

Tissue grasping and handling is essential in surgery, yet excessive clamping force can cause tissue injury [1]. With conventional forceps, the applied force is regulated manually by the surgeon and can vary with the operator input, grasped object size, and jaw configuration. Although surgeons may infer the interaction force from visual and haptic cues, these cues provide only indirect information about the actual tissue interaction force, particularly in minimally invasive procedures [2].

Micro force-sensing instruments quantify interaction forces [3] but add cost, complexity, and sterilization burden. Passive force-limiting forceps [1] offer a simpler alternative, but regulate the clamping force to a fixed, non-adjustable level. Constant-force compliant mechanisms enable the force level to be adjusted [4]. However, in a forceps designed on this principle [5], the grasping torque is determined by the instantaneous handle angle. Maintaining a constant force therefore requires the surgeon to continuously hold the handle at a fixed angle.

This paper presents an Adjustable Constant-Force (ACF) surgical forceps based on a passive compliant mechanism that combines positive and negative stiffness to create a Quasi-Zero-Stiffness (QZS) regime. The proposed design removes the handle angle as a continuous force-control variable. Instead, the desired force is mechanically preset by an adjustable end-stop. The surgeon closes the handle against the stop, after which the mechanism maintains an approximately constant clamping force over the jaw opening range. The mechanism is fully passive, electronics-free, and inexpensive to manufacture.

MATERIALS AND METHODS

A. Adjustable Constant-Force Mechanism

As shown in Fig. 1(a), the ACF mechanism comprises a nonlinear Negative-Stiffness Mechanism (NSM), a positive-stiffness spring, a slider, a handle, an adjustment nut, and a tendon-driven jaw assembly. The positive spring ($K_p > 0$) provides a restoring force that increases linearly with slider displacement x. The NSM is realized by a pre-compressed inclined spring. It exhibits negative incremental stiffness $k_{\mathrm{NSM}}(x) < 0$. The positive spring and the NSM act in parallel on the slider, such that their stiffness contributions combine to give the equivalent stiffness:

$$k_{\mathrm{eq}}(x) = \frac{\mathrm{d}F_s}{\mathrm{d}x} = K_p + k_{\mathrm{NSM}}(x) \approx 0, \qquad (1)$$

where F_s is the net force transmitted from the slider to the tendon. The parameters are chosen so that the two contributions nearly cancel over the operational stroke. This is the QZS condition, which makes F_s approximately constant. A non-stretchable tendon, driven by the slider rather than directly by the handle, transmits F_s to the movable jaw through a pulley. With a constant pulley radius, an approximately constant F_s therefore produces an approximately constant clamping torque. The tendon displacement determines the jaw motion, while contact with the grasped object sets the jaw opening angle and, correspondingly, the slider position x. Since F_s is approximately insensitive to x within the QZS region, the clamping torque remains nearly constant across the operational range of jaw openings.

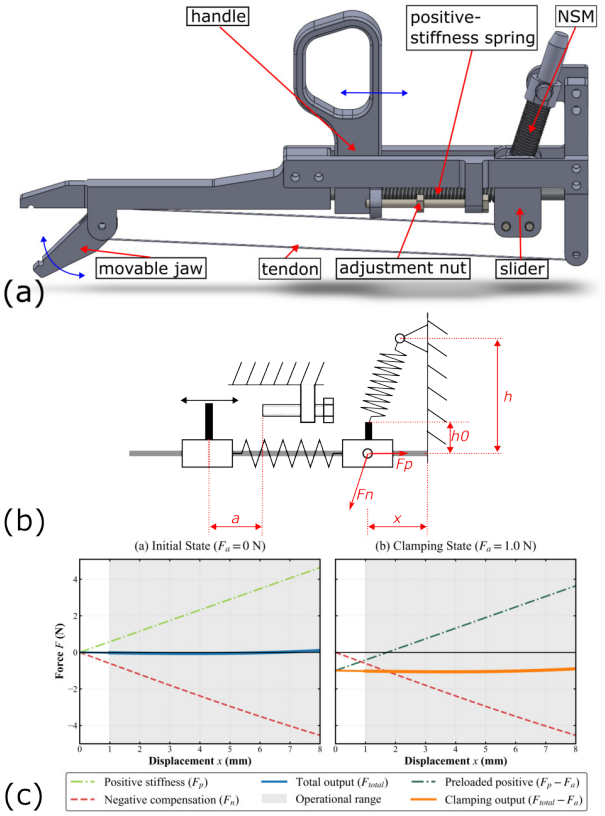


Fig. 1: Overview of the ACF forceps: (a) CAD model with main components labelled. (b) Kinematic schematic defining slider displacement x, handle stroke a, anchor height h, offset h_0, and spring forces F_p, F_n. (c) Force superposition at $F_a = 0$ (left) and at a representative nonzero setting ($F_a \equiv K_p\, a > 0$, right): F_p and F_n cancel within the operational stroke, leaving a net slider output equal to F_a.

B. Modeling and Adjustment

In Fig. 1(b), the positive spring (K_p) links the slider to a fixed support and the nut sets the handle stroke a. The NSM spring (K_n, free length L_0) is pre-compressed between a wall anchor at height h and a rigid link of height h_0 on the slider. Let x denote the slider displacement from its initial position. The positive spring then exerts $F_p(x, a) = K_p(x + a)$. The horizontal reaction force of the inclined

This work was supported by Startup Grant of the Danish Institute of Advanced Study (DIAS) and VP-GRACE project.

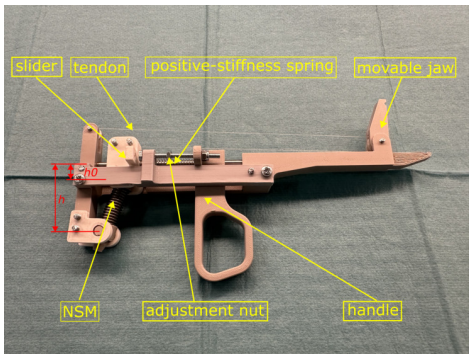


Fig. 2: Prototype of the ACF forceps.

spring is

$$F_n(x) = K_n\, x \left(\frac{L_0}{\sqrt{x^2 + (h - h_0)^2}} - 1 \right), \qquad (2)$$

the net tendon force is

$$F_s(x, a) = K_p(x+a) - F_n(x) = K_p\, a + \left[K_p\, x - F_n(x) \right]. \quad (3)$$

The first term $K_p a$ is set by the nut's position. The bracketed term collects all x-dependent contributions, representing the effect of grasped object size on the output force. The parameters $(K_p, K_n, L_0, h - h_0)$ are selected such that this term remains close to zero over the operational stroke. The output force then reduces to

$$F_s(x, a) \approx K_p\, a \equiv F_a. \qquad (4)$$

Fig. 1(c) shows this force superposition: when $F_a = 0$, the positive- and negative-stiffness contributions approximately cancel, whereas for nonzero a, their superposition yields a net output force of F_a. The anchor height h is chosen to minimize the variation of F_s over the operational stroke.

C. Prototype

The prototype is shown in Fig. 2. The NSM is formed by a standard compression spring ($L_0 = 68$ mm, $K_n = 0.44$ N/mm), pre-compressed between a wall anchor at $h = 37.6$ mm and a link of height $h_0 = 9$ mm on the slider. The spring tilts as the slider moves forward, so its horizontal reaction $F_n(x)$ in Eq. (2) grows nonlinearly, giving the negative stiffness. A 1-mm slider offset avoids the singular configuration at $x = 0$, where the inclined spring is perpendicular to the shaft and its horizontal reaction vanishes. The resulting operational stroke is $x \in [1, 8]$ mm. The positive spring ($K_p = 0.58$ N/mm) is connected in series between the slider and the handle. The nut constrains the handle stroke to a. The tendon drives the movable jaw over a pulley of radius $R = 5$ mm, giving $\theta = \frac{s}{2\pi R} \times 360^\circ \approx 80^\circ$ over the $s = 7$ mm stroke.

EXPERIMENTS AND RESULTS

The jaw output torque was measured at five jaw opening angles under three preset clamping-force levels. As shown in Fig. 3(a) and (b), the forceps and a force/torque sensor (FTS-Gamma 047, SCHUNK, Germany; in-plane resolution: 0.0063 N) were rigidly mounted on a 3D-printed fixture. A rod was attached to the center of the sensor's sensing surface. During each measurement, the movable jaw was brought into contact with the rod, allowing the sensor to measure the clamping force. The fixed jaw was positioned at five distinct angles relative to the movable jaw: 15°, 30°, 45°, 60°, and 75° with a 25 mm moment arm. The adjustment nut was set to 8 mm, 12 mm, and 16 mm to prescribe three constant clamping-force levels, with six measurements performed for each angle–setting combination.

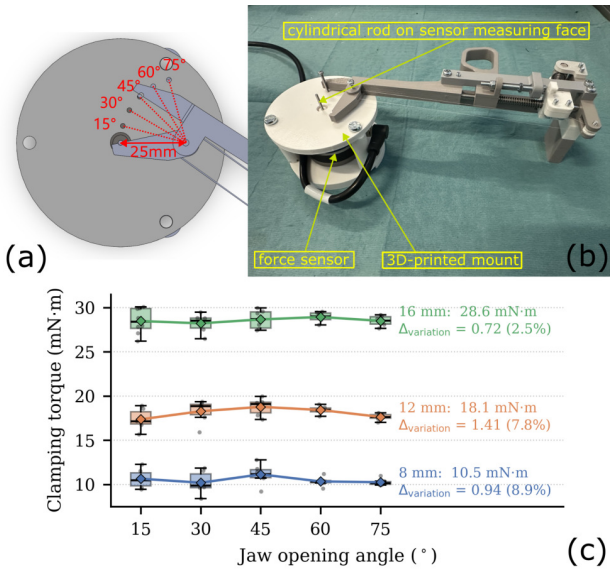


Fig. 3: (a) The fixed jaw was positioned at five different angles relative to the movable jaw with a 25 mm moment arm. (b) Test setup with the force/torque sensor. (c) Clamping torque vs. jaw opening angle at the three nut settings. Boxes show the interquartile range and median, diamonds the mean torque, and $\Delta_{\text{variation}}$ is the max–min variation of the mean torque across jaw opening angles.

Fig. 3(c) shows the distribution of the measured clamping torque at each jaw opening angle, together with the corresponding mean values. The torque range, defined as the difference between the maximum and minimum mean torque across the five jaw opening angles, is 0.94, 1.41, and 0.72 mN·m for the 8, 12, and 16 mm settings, respectively. These ranges correspond to 8.9%, 7.8%, and 2.5% of the respective overall mean torques, indicating that the clamping torque is largely insensitive to the jaw opening angle over the tested range.

CONCLUSIONS AND DISCUSSION

We presented an ACF surgical forceps that combines positive and negative stiffness to maintain a nearly constant clamping force across different object sizes and jaw opening angles. Experiments across five jaw opening angles and three force settings showed that the variation in average clamping torque remained below 9%, confirming the constant-force behavior over the tested range. Future work will calibrate the nut setting to allow direct force-based adjustment.

REFERENCES

[1] Y. Maeda *et al.*, "Development of novel force-limiting grasping forceps with a simple mechanism," *European Journal of Cardio-Thoracic Surgery*, vol. 54, no. 6, pp. 1004–1011, 2018.

[2] E. A. M. Heijnsdijk *et al.*, "The influence of force feedback and visual feedback in grasping tissue laparoscopically," *Surgical Endoscopy*, vol. 18, no. 6, pp. 980–985, 2004.

[3] U. Kim, D.-H. Lee, W. J. Yoon, B. Hannaford, and H. R. Choi, "Force sensor integrated surgical forceps for minimally invasive robotic surgery," *IEEE Trans. on Robot.*, vol. 31, no. 5, pp. 1214–1224, 2015.

[4] Z. Cheng, T. R. Savarimuthu, S. Foong, and U.-X. Tan, "Design of adjustable constant force/torque mechanisms for medical applications," *Journal of Mechanisms and Robotics*, vol. 15, no. 2, p. 025001, 2023.

[5] Z. Cheng, "Forceps with direct torque control," in *Proc. Conf. on New Technologies for Computer/Robot Assisted Surgery (CRAS)*, 2023.

An Open-Source Environment for Autonomous Robotic Colonoscopy Navigation

Guanglin Ji[1], Martina Finocchiaro[1], Kenny Erleben[1], Hang Yin[1]

[1]*Department of Computer Science, University of Copenhagen, Copenhagen, Denmark.*

***Abstract*— Autonomous robotic colonoscopy requires navigation in long, deformable, and visually ambiguous anatomy. Existing simulators often use rigid tubes or fully actuated capsule robots, limiting their ability to study contact-rich navigation of clinically relevant flexible endoscopes. This paper presents ARCGym, an Isaac Sim/Isaac Lab-based tutorial environment for robotic colonoscopy simulation. ARCGym provides a Gymnasium-compatible interface, deformable finite-element colon models, egocentric RGB-D sensing, and multiple robot abstractions, including capsule, magnetic, and articulated flexible endoscopes. The environment supports teleoperation and future policy training, enabling reproducible studies of robot–tissue interaction, observation design, action parameterization, and contact stability.**

INTRODUCTION

Colonoscopy is widely used for colorectal cancer screening, but manual navigation remains operator-dependent due to complex anatomy, tissue deformability, frictional contact, and limited endoscopic visibility. Recent reviews have highlighted robotic, self-propelled, and self-steerable colonoscopes as alternatives to conventional manual insertion [1]. However, their development requires simulators that capture realistic robot–colon interaction.

Existing simulators often use rigid tubes, idealized centerlines, or fully actuated capsule-like robots. These abstractions are useful but insufficient for long flexible endoscopes, whose shafts continuously interact with the bowel wall and may induce deformation, friction, looping, or jamming, especially in the sigmoid colon. Although autonomous colonoscopy has been explored using magnetic actuation, lumen-center detection, and learning-based control [2], [3], [4], contact-rich simulation remains necessary because proximal insertion may either advance the tip or merely store elastic energy in the robot–tissue system.

We present ARCGym, an Isaac Sim/Isaac Lab-based environment for robotic colonoscopy simulation as shown in Fig. 1. ARCGym integrates deformable colon models, egocentric camera sensing, multiple robot embodiments, teleoperation support, and a Gymnasium-compatible interface. The open-source code will be available when the abstract is accepted.

MATERIALS AND METHODS

I. Software Setup and Launching

ARCGym is implemented on top of NVIDIA Isaac Sim/Isaac Lab, PhysX, Gymnasium, and PyTorch. The environment uses PhysX for rigid-body articulation and deformable-body simulation, exposes a Gymnasium-compatible API, and supports Isaac Lab tiled RGB-D cameras with a tip-mounted light. Although this paper focuses on teleoperation-mode demonstrations, the same interface is compatible with Stable-Baselines3 for future policy learning. The main entry point launches the Isaac application, enables camera rendering, registers the ARCGym environment, and creates folders for trajectories, checkpoints, and visual outputs. For teleoperation and visualization with N parallel colon environments, ARCGym can be launched as: `python arc_env_rl.py --num_envs N`. For future policy learning, the same entry point can be used with parallel environments. For example, PPO training can be launched as: `python arc_env_rl.py --train --num_envs N`.

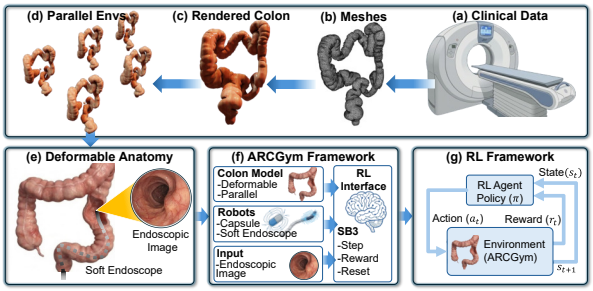


Fig. 1: Overview of ARCGym. Clinical colon data are converted into anatomical meshes and rendered as deformable Isaac Sim environments. ARCGym supports parallel simulation, multiple robot embodiments, egocentric endoscopic observations, and a Gymnasium-compatible interface for teleoperation, data collection, and future learning-based control.

II. Selecting a Robot Model

Robot models are selected through the `robot_type` field in the robot configuration. ARCGym uses a robot factory abstraction, which instantiates the corresponding robot class while keeping the same environment interface. The currently supported options are:

```
robot_type = "capsule"
robot_type = "magnetic_endoscope"
robot_type = "soft_endoscope"
```

For this paper, two robot types are most relevant. The `magnetic_endoscope` model represents a passively articulated robot driven by a distal 6-DoF command. The `soft_endoscope` model represents a proximally inserted flexible endoscope with feed, bending, twist, passive shaft compliance, and a tip-mounted RGB-D camera.

III. Configuring the Colon Model

For teleoperation demonstrations, the colon is configured as a deformable FEM mesh. The key issue is not the rigid/deformable switch, but the selection of consistent material, damping, and contact parameters. A default setting is:

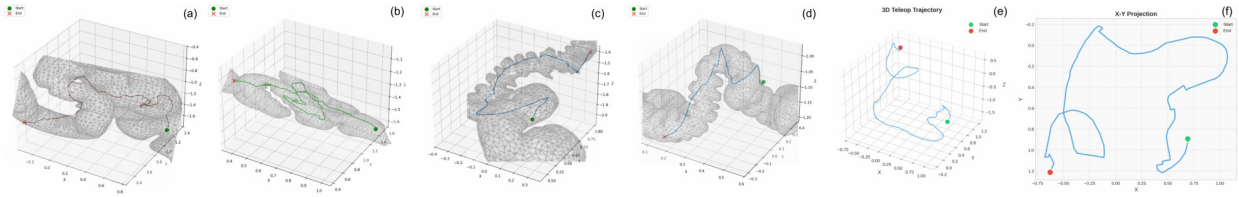


Fig. 2: Representative navigation trajectories generated in ARCGym. Panels (a)–(d) show teleoperated robot trajectories within different colon anatomies, where green and red markers denote the start and end points, respectively. Panels (e)–(f) summarize the corresponding 3D tip trajectory and its planar projection, illustrating the feasibility of trajectory recording and visualization during teleoperated colonoscopy simulation.

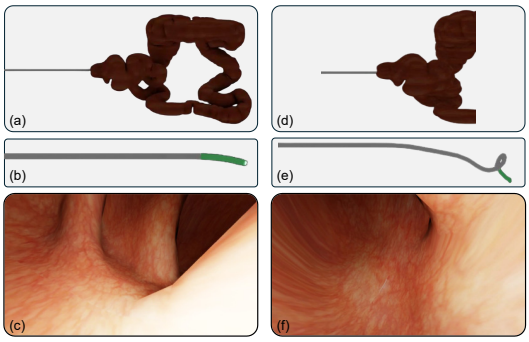


Fig. 3: Teleoperation snapshots in ARCGym. The figure compares the initial configuration and the navigated sigmoid state using global colon views, endoscope configurations, and local endoscopic camera observations.

```
colon_mass = 1.5
youngs_modulus = 1.0e5
poissons_ratio = 0.49
elasticity_damping = 50.0
```

These parameters should be tuned jointly with robot mass and passive joint impedance. High stiffness with low damping may cause unrealistic rebound, whereas overly soft tissue or inappropriate contact offsets may cause excessive deformation, wall penetration, or unstable contact.

IV. Configuring the Flexible Endoscope

For the proximally inserted endoscope, the most important parameters are the number of links, link geometry, link mass, passive shaft stiffness/damping, distal bending stiffness/damping, twist stiffness/damping, and maximum feed velocity. A default configuration is:

```
robot_type = "soft_endoscope"/"magnetic_endoscope"
passive_stiffness = 500.0
passive_damping = 800.0
active_stiffness = 5000.0
active_damping = 1000.0
twist_stiffness = 5000.0
twist_damping = 1000.0
```

The passive shaft parameters are particularly important for colonoscopy simulation. If the passive shaft is too soft, proximal insertion can create an S-loop and the tip may stop advancing. If it is too stiff, the robot cannot conform to curved anatomy and may push against the colon wall. Therefore, teleoperation is used as a first-stage software sanity check before any policy learning.

RESULTS

Teleoperation demonstrations validate ARCGym from both external and egocentric viewpoints. As shown in Fig. 3, the simulator records the initial rectal insertion state and the navigated sigmoid state, including the global colon configuration, the endoscope shape, and the corresponding local endoscopic camera view. This paired visualization allows users to inspect whether camera-level lumen feedback is consistent with the underlying robot–colon interaction.

Fig. 2 further shows that ARCGym can record navigation trajectories across different colon anatomies, including full 3D tip paths and their planar projections. For the `soft_endoscope`, teleoperation uses proximal feed/withdrawal, distal bending, and axial twist; for the `capsule` and `magnetic_endoscope`, control is represented as a 6-DoF distal velocity command. These demonstrations provide a practical sanity check before policy learning: forward insertion should produce measurable tip advancement rather than only shaft bending, wall sticking, or tissue loading.

CONCLUSIONS AND DISCUSSION

ArcGym provides a practical prototype for colonoscopy robot training using Isaac Sim, Isaac Lab, and reinforcement learning. The platform supports both manual teleoperation and automated PPO training. ARCGym can serve as a unified navigation-strategy learning platform for capsule robots, magnetically driven flexible endoscopes, and cable-driven endoscopes, while also enabling the generation of synthetic visual, state, and action data for future world-model and vision–language–action (VLA) research.

REFERENCES

[1] C. Winters, V. Subramanian, and P. Valdastri, "Robotic, self-propelled, self-steerable, and disposable colonoscopes: Reality or pipe dream? a state of the art review," *World Journal of Gastroenterology*, vol. 28, no. 35, p. 5093, 2022.

[2] P. R. Slawinski, A. Z. Taddese, K. B. Musto, K. L. Obstein, and P. Valdastri, "Autonomous retroflexion of a magnetic flexible endoscope," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1352–1359, 2017.

[3] V. S. Nguyen, B. Hwang, B. Kim, and J. H. Jung, "An end-to-end learning-based control signal prediction for autonomous robotic colonoscopy," *IEEE Access*, vol. 12, pp. 1280–1290, 2023.

[4] Y. Li, W. Y. Ng, Y. Sun, Y. Huang, J. Li, P. W. Y. Chiu, and Z. Li, "Colon lumen center detection enables autonomous navigation of an electromagnetically actuated soft-tethered colonoscope," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–15, 2024.

Intracardiac Soft Robotic Design for CPR Application

Zhihao Ma, Linxiaohai Ning, Xinxin Li, Tim Harris, and Zion Tse

Queen Mary University of London, E1 4NS London, U.K.

INTRODUCTION

After cardiac arrest, cardiopulmonary resuscitation (CPR) provides temporary mechanical circulatory support until defibrillation or other definitive treatment is available. Its effectiveness depends on early application and compression quality. The compressive force is directed at the heart to provide circulation but also affects other organs, such as the sternum, ribs, lungs and abdominal structures, with potential organ damage. Because this load path includes compliant thoracic structures in a variable orientation, similar external inputs can produce variable cardiac loading across compression locations. During resuscitation, coronary perfusion pressure (CPP), the difference between aortic diastolic pressure and left ventricular end-diastolic pressure, determines myocardial perfusion and higher values > 15 mmHg are required to obtain return of spontaneous circulation (ROSC), although achieving a given CPP does not guarantee survival [1]. Currently, closed-chest cardiac massage is the first-line approach, but it compresses the heart indirectly through the thoracic wall and can be limited by chest-wall mechanics, rescuer fatigue, and interruptions in compressions. Existing CPR-assist devices can standardise external compression, but they still transmit force indirectly through the chest wall and may impose substantial loads on thoracic and abdominal tissues.

Applying force directly to the heart's surface may reduce this transmission limitation by bypassing part of the compliant thoracic load path. Open-chest cardiac massage (OCCM) can provide such compression in selected emergency contexts, notable traumatic cardiac arrest with direct myocardial injury. However, OCCM is invasive, requires surgical access and sterile technique, and may introduce tissue injury and infection [2]. Previous minimally invasive direct cardiac massage concepts have suggested that direct compression through a limited thoracic incision is technically feasible, but reliable placement, procedural safety, and clinical efficacy remain unresolved barriers [3].

Soft robotics is attractive for direct cardiac compression because compliant fluidic structures can conform to curved biological surfaces, distribute contact pressure, and undergo large deformation with lower local stress than rigid tools [4]. Eversion robots are especially useful for deployment because they grow from the advancing tip rather than sliding along their full length, which can reduce friction in confined spaces [5]. In this prototype, the eversion robot creates and positions the pathway, whereas the catheter balloon supplies local normal compression. A commercial urinary catheter is used only as a temporary proof-of-concept balloon and will be replaced by a customised biocompatible compression chamber in future designs. This paper reports an early bench-top proof of concept in a laboratory chest model with a silicon heart phantom.

MATERIALS AND METHODS

The prototype consists of two soft eversion robots, two catheter-based inflatable balloon modules, and a pneumatic actuation setup (Fig. 1). The eversion robots are polyethylene-film tube-like structures stored in a compact container before actuation. When pneumatic pressure is applied, each robot everts from the inside out and extends toward the target region. The working configuration of the delivery and compression module is shown in Fig. 2. For deployment testing, the prototype was introduced through an approximately 6 mm $\times$ 20 mm access opening at the fourth intercostal space of the laboratory chest model.

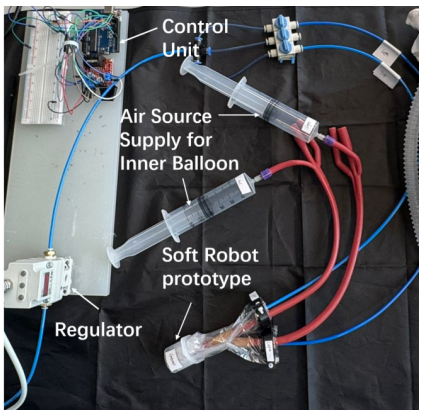


Fig. 1: Prototype and pneumatic setup, including the controller, regulator, air source, soft eversion robots, and syringe-based supply for the inner catheter balloons.

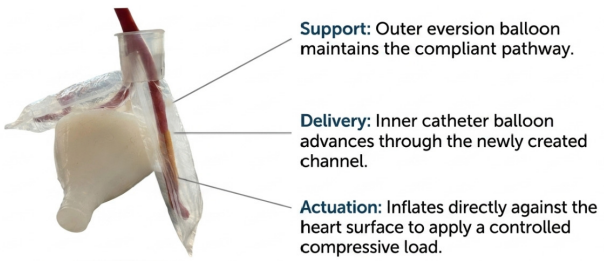


Fig. 2: Working configuration of the soft eversion delivery pathway and inner catheter-balloon compression module.

The eversion robot creates a temporary compliant access pathway and guides the catheter balloon module toward the silicon heart phantom. After reaching the target region, the catheter balloon is inflated and deflated to apply cyclic

local compression. An electromagnetic tracking marker near the target provided displacement relative to the pre-test baseline. Output force was measured with a load cell during single catheter-balloon compressions. Phantom stroke volume was defined as the expelled phantom volume per maximum compression. Successful deployment was defined as eversion through the intercostal access opening, positioning of the catheter balloon near the target phantom region, and sustained cyclic actuation without loss of contact.

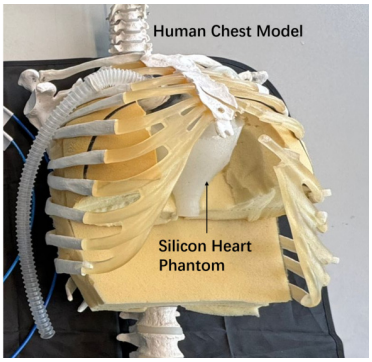


Fig. 3: Laboratory chest model and silicon heart phantom used for bench-top deployment and compression tests.

CONCLUSIONS AND DISCUSSION

An eversion input pressure of approximately 4 psi deployed the outer eversion balloons through the model intercostal access space and guided them toward the target region. The inner catheter balloons were then advanced within the eversion pathways and actuated using a 60 mL syringe. The load-cell measurement showed that a single catheter-balloon compression generated approximately 2–3 N.

Electromagnetic tracking of a marker attached to the phantom measured local displacement during the nominal cyclic-compression test. This measurement is separate from the maximum-compression volume benchmark reported in Table I. The trace is relative to the pre-test marker position and contains two mechanical components. The initial rise reflects gross loading from outer eversion-balloon deployment, reaching 16.3 mm. After positioning, the inner catheter balloon produced approximately 10 mm cyclic oscillations for 60 s. These cycles occurred around a deployment-induced offset, suggesting maintained contact or preload. Small cycle-to-cycle variations are expected for manually supplied syringe input and a compliant phantom.

As a separate maximum-compression benchmark, phantom volume output was measured as expelled phantom volume per compression and converted to $EF_{\text{phantom}} = \text{SV}/\text{EDV} \times 100\%$, where EDV = 56 mL. Unlike Fig. 4, Table I reports upper-bound tests rather than nominal cyclic compression. Manual direct compression produced 23.3 mL per beat and a phantom-derived EF of 42.0%. The catheter-balloon configuration produced 9–11 mL per beat, corresponding to 16.1–19.6% phantom-derived EF, or approximately 39–47% of the manual maximum-compression volume output. These results support the overall architecture of separating soft eversion-based delivery from inflatable compression actuation.

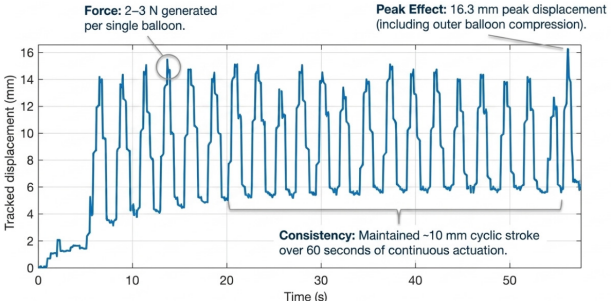


Fig. 4: Tracked local phantom displacement during nominal cyclic compression. The initial rise is produced by outer-balloon deployment. The subsequent approximately 10 mm oscillations correspond to inner catheter-balloon actuation over 60 s. These displacement data are distinct from the maximum-compression volume data in Table I.

TABLE I: Maximum-compression volume output in the silicon heart phantom. SV is the expelled phantom volume per compression, and EF is calculated using EDV = 56 mL.

Method	SV (mL/beat)	EF (%)
Hand	23.3	42.0
Catheter balloon	9–11	16.1–19.6

Overall, these results support a preliminary mechanical proof of concept for direct phantom compression using a soft eversion delivery pathway and temporary catheter-balloon actuator. The laboratory system was introduced through an approximately 6 mm × 20 mm access opening at the fourth intercostal space of the chest model and achieved positioning near the phantom, measurable force, nominal cyclic displacement, and maximum-compression phantom volume output. Future work should replace the temporary urinary catheter with a customised biocompatible compression chamber with defined contact area, inflation volume, and pressure limits. Subsequent studies should quantify repeatability under automated actuation, measure pressure–flow output, and evaluate tissue interaction, access-site trauma, infection risk, and sterile thoracic access.

REFERENCES

[1] P. A. Meaney, B. J. Bobrow, M. E. Mancini, J. Christenson, A. R. de Caen, F. Bhanji, B. S. Abella, M. E. Kleinman, D. P. Edelson, R. A. Berg, T. P. Aufderheide, V. Menon, and M. Leary, "Cardiopulmonary resuscitation quality: Improving cardiac resuscitation outcomes both inside and outside the hospital: A consensus statement from the american heart association," *Circulation*, vol. 128, no. 4, pp. 417–435, 2013.

[2] R. S. D'Souza, S. Sharma, S. A. Collier, and L. Law, "Open heart massage," in *StatPearls*. Treasure Island, FL: StatPearls Publishing, 2024.

[3] C. A. Eynon, "Minimally invasive direct cardiac massage and defibrillation," *Resuscitation*, vol. 47, no. 3, pp. 325–328, 2000.

[4] E. T. Roche, M. A. Horvath, I. Wamala, A. Alazmani, S.-E. Song, W. Whyte, Z. Machaidze, C. J. Payne, J. C. Weaver, G. Fishbein, J. Kuebler, N. V. Vasilyev, D. J. Mooney, F. A. Pigula, and C. J. Walsh, "Soft robotic sleeve supports heart function," *Science Translational Medicine*, vol. 9, no. 373, p. eaaf3925, 2017.

[5] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, "A soft robot that navigates its environment through growth," *Science Robotics*, vol. 2, no. 8, p. eaan3028, 2017.

A Re-positionable and Axially Re-orientable Pivot-Joint Steering Mechanism for Tip-Everting Soft Growing Robots

Tianchen Ji[1], Diyang Wu[1], Zheyuan Bi[1], and Lin Cao[1,*]

[1]*School of Electrical and Electronic Engineering, The University of Sheffield, The United Kingdom*
[*]*Corresponding author. Email: l.cao@sheffield.ac.uk*

INTRODUCTION

Endoluminal medical navigation, such as colonoscopy and bronchoscopy, requires long, slender robots that can access tortuous anatomical pathways while minimizing tissue interaction forces. Conventional flexible endoscopes are effective but are advanced by pushing from the base, which can create discomfort, looping, and tissue stress in highly curved anatomy [1], [2]. Soft growing robots (SGRs), or vine robots, offer an alternative strategy: they advance by tip eversion, so the deployed body remains nearly stationary relative to the environment, reducing friction during navigation [3], [4], [5]. This makes SGRs promising for minimally invasive procedures and other confined-space inspection tasks [6].

However, steering a continuously growing SGR remains difficult. Distributed pneumatic muscles or tendon-driven bodies can steer the robot, but often bend the entire body or require actuators along the robot length [7], [8]. Internal steering, hydraulic skeletons, magnetic skins, and inchworm-inspired mechanisms improve specific capabilities, but can introduce fixed steering regions, external-field requirements, multiple actuators, or task-specific hardware [9], [10], [5], [2]. A compact steering unit that can be relocated and reused during growth is therefore desirable, especially for medical-scale systems where lumen space is limited.

Here, we extend our previous pivot-joint SGR [11] toward spatial locomotion. The robot uses a tendon-driven pivot joint for localized bending, a pressure-tunable circumferential bladder for frictional coupling, and a torque-transmitting sheath for axial relocation and steering-plane reorientation. This creates a compact "steer-grow-reposition-rotate" strategy, allowing a single internal steering module to be repeatedly reused for planar and spatial navigation without distal motors or distributed steering actuators.

MATERIALS AND METHODS

A. Pivot-Joint Relocation Along the Robot Body

For sustained steering, the internal joint must be moved toward the distal region as the robot grows. Figure 1 shows two relocation modes. When the bladder is deflated, the pivot-joint holder is weakly coupled to the inverted membrane, allowing the torque sheath to directly push the joint forward (Fig. 1(a,b)). When the bladder is inflated, the bladder increases the normal force on the inverted membrane and couples the joint to the moving membrane (Fig. 1(d,e)); the joint is then dragged forward during eversion (Fig. 1(c)). The first mode enables deliberate repositioning, while the second mode synchronizes joint transport with growth. Together, they allow the same compact joint to be reused for repeated "steer–grow–reposition" operation.

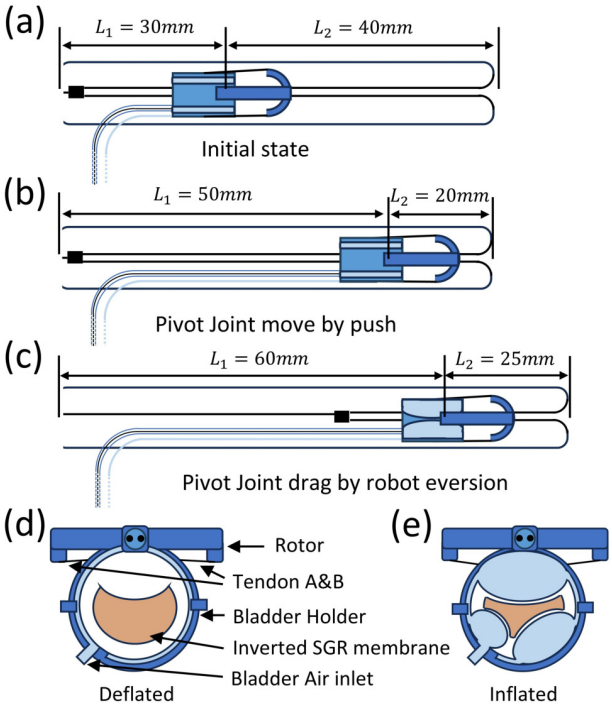


Fig. 1: Pivot-joint relocation and bladder coupling. (**a–b**) Push-based relocation with the bladder deflated. (**c**) Eversion-assisted relocation with the bladder inflated. (**d–e**) Bladder-based frictional coupling between the joint and the inverted membrane.

B. Steering and Steering-Plane Reorientation

Steering is generated by converting tendon tension into local rotor rotation. The rotor lever arm is $r = \sqrt{l_1^2 + l_2^2}$. For a tendon force F, the effective normal force is $F_n = F \sin\theta_1$, producing a rotor moment $M = F_n r$. The angle θ_1 depends on the rotor angle and tendon-routing offsets l_3 and l_4 (Fig. 2(a,b)). After the rotor and bladder holder contacts the robot wall, further tendon pulling locally compresses the inflated membrane and creates a bending constraint at the

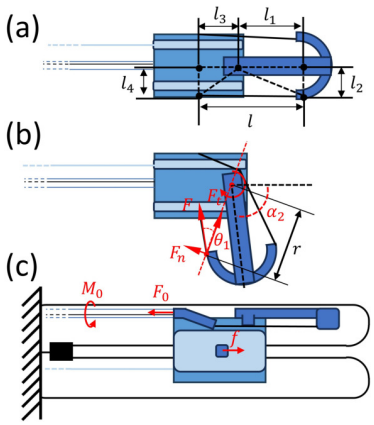


Fig. 2: Steering and steering-plane reorientation. (**a**) Rotor and tendon-routing geometry. (**b**) Tendon-driven rotor rotation for localized steering. (**c**) Axial rotation of the pivot-joint module through the torque sheath.

pivot-joint location. This redirects subsequent tip eversion without bending the entire robot body.

The steering plane is changed by transmitting torque M_0 through the sheath (Fig. 2(c)). This rotates the whole pivot-joint module about the robot axis, allowing the same tendon-driven joint to steer in a different plane. Thus, spatial steering is achieved by alternating localized bending and sheath-driven axial reorientation.

C. Steering Force and Bladder Friction

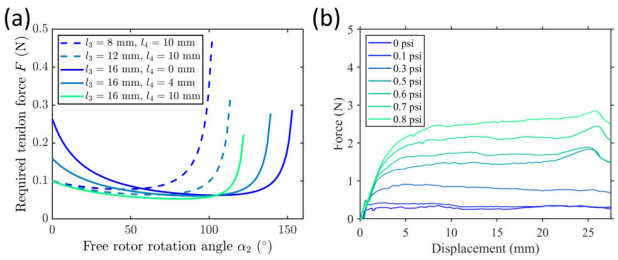


Fig. 3: Steering-force and bladder-friction analysis. (**a**) Required tendon force for generating 1 $N{\cdot}mm$ rotor moment under varying tendon-routing geometries. (**b**) Measured friction force generated by the circumferential bladder under different bladder pressures.

Figure 3(a) indicates that tendon-routing geometry determines how efficiently tendon tension is converted into rotor moment. Increasing the horizontal offset l_3 reduces the required force and delays the rapid force increase at large rotor angles, while increasing the vertical offset l_4 lowers the force demand in the small-angle range but shifts the sharp rise to smaller rotor angles. These trends provide practical guidance for selecting routing geometry to reduce actuation load within the desired steering range.

The circumferential bladder provides the clutch-like coupling required for joint holding and relocation. Its available friction can be approximated as $f_B = \mu P_B A_B$, where P_B is bladder pressure, A_B is contact area, and μ is the friction coefficient. As shown in Fig. 3(b), the friction force remains low when the bladder is unpressurised, enabling sliding during push-based relocation. Increasing bladder pressure substantially increases the holding force; at 0.8 psi, the bladder provides more than 2.5 N of friction over most of the tested displacement range, which is sufficient for high-friction membrane coupling during steering and eversion-assisted transport.

D. Navigation Demonstrations

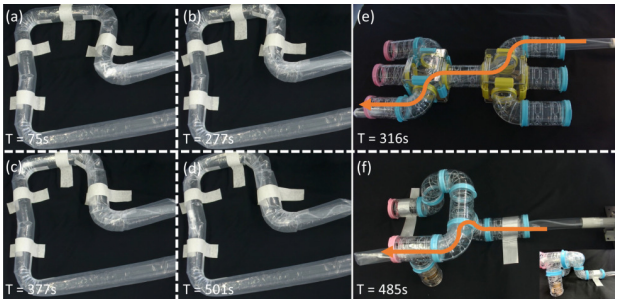


Fig. 4: Navigation demonstrations. (**a–d**) Navigation through a colon-like tube using repeated steering, growth, and repositioning. (**e**) Planar cluttered pipe navigation. (**f**) Spatial pipe navigation enabled by steering-plane reorientation.

The robot was tested in medical-relevant and cluttered tube environments. In the colon-like tube, the robot navigated multiple curved sections by repeating localized pivot-joint steering, tip eversion, and downstream joint repositioning (Fig. 4(a–d)). In the planar pipe environment, the same sequence guided the robot through multiple bends (Fig. 4(e)). In the spatial pipe environment, sheath-driven axial reorientation changed the steering plane before subsequent growth, enabling non-planar navigation (Fig. 4(f)).

CONCLUSIONS AND DISCUSSION

This work presents a spatially steerable pivot-joint SGR for confined and endoluminal navigation. The proposed mechanism combines localized tendon-driven steering, pressure-tunable bladder coupling, direct joint relocation, and steering-plane reorientation. Compared with distributed actuation or fixed steering mechanisms, a single compact internal joint can be repeatedly reused during growth.

The results show that tendon-routing geometry can reduce steering force, while the bladder provides a tunable friction interface for sliding, holding, and membrane-coupled transport. Demonstrations in colon-like and cluttered tube environments show repeated ”steer–grow–reposition” operation and spatial steering through axial joint rotation. These capabilities are relevant to future medical SGRs, where small diameter, low tissue interaction, and reusable distal steering are critical.

Future work will focus on closed-loop control, miniaturization, improved bladder durability, and integration of imaging or therapeutic tool channels for endoluminal applications.

REFERENCES

[1] Y. Li, C. Guo, W. Xin, T. Pan, W. Li, P. W. Y. Chiu, and Z. Li, "Design and preliminary evaluation of an electromagnetically actuated soft-tethered colonoscope," *IEEE Transactions on Medical Robotics and Bionics*, vol. 3, no. 2, pp. 402–413, 2021.

[2] A. Giri, C. Girerd, J. Cervera-Torralba, M. T. Tolley, and T. K. Morimoto, "InchIGRAB: An inchworm-inspired guided retraction and bending device for vine robots during colonoscopy," *IEEE/ASME Transactions on Mechatronics*, 2025, early access.

[3] E. W. Hawkes, L. H. Blumenschein, J. D. Greer, and A. M. Okamura, "A soft robot that navigates its environment through growth," *Science Robotics*, vol. 2, no. 8, p. eaan3028, 2017.

[4] J. D. Greer, L. H. Blumenschein, A. M. Okamura, and E. W. Hawkes, "Obstacle-aided navigation of a soft growing robot," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018, pp. 4165–4172.

[5] J. Davy, N. J. Greenidge, S. Kim, L. J. Tinsley, P. Lloyd, J. H. Chandler, R. A. Harris, T. K. Morimoto, and P. Valdastri, "Vine robots with magnetic skin for surgical navigations," *IEEE Robotics and Automation Letters*, vol. 9, no. 8, pp. 6888–6895, 2024.

[6] A. Oyejide, F. Stroppa, and M. Sarac, "Miniaturized soft growing robots for minimally invasive surgeries: Challenges and opportunities," *Progress in Biomedical Engineering*, vol. 7, p. 033001, 2025.

[7] J. D. Greer, T. K. Morimoto, A. M. Okamura, and E. W. Hawkes, "A soft, steerable continuum robot that grows via tip extension," *Soft Robotics*, vol. 6, no. 1, pp. 95–108, 2019.

[8] S. Wang, R. Zhang, D. A. Haggerty, N. D. Naclerio, and E. W. Hawkes, "A dexterous tip-extending robot with variable-length shape-locking," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 9035–9041.

[9] D. A. Haggerty, N. D. Naclerio, and E. W. Hawkes, "Hybrid vine robot with internal steering-reeling mechanism enhances system-level capabilities," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5437–5444, 2021.

[10] T. Takahashi, K. Tadakuma, M. Watanabe, E. Takane, N. Hookabe, H. Kajiahara, T. Yamasaki, M. Konyo, and S. Tadokoro, "Eversion robotic mechanism with hydraulic skeleton to realize steering function," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5413–5420, 2021.

[11] T. Ji, Z. Bi, and L. Cao, "A pivot joint steering mechanism for tip-everting soft growing robots," *Frontiers in Robotics and AI*, vol. 12, p. 1627116, 2025.

Author Index

Abdo, Nicolas — 22
Abdulali, Arsen — 108
Adagolodjo, Yinoussa — 29
Agarwal, Agrima — 50
Ahmed, Aminul — 94
Akanyeti, Otar — 46, 68
Akpo, Eugenia — 11
Al Haimus, Muhie — 75
Ali, Sharib — 64
Althoefer, Kaspar — 75, 108
Amat, Josep — 57
Anchal, Sahil — 61
Andrew, Bailey — 62
Angelmahr, Martin — 108
Argent Masa, Carlos — 56
Artinian, Azad — 27
Arusi, Shani — 37
Avery, James — 48, 105, 107
Azad, Bita — 9
Bacchetti, Alistair — 82, 97
Bano, Edwige — 67
Barata, Beatriz — 102
Barrett, Tristan — 47
Basoli, Valentina — 38
Beckers, Wim-Alexander — 102
Bellini, Elisa — 35
Bennett, Johan — 43
Bennett, Rory — 47
Bennett, Wayne — 28
Bergeles, Christos — 11, 28, 94
Bernabei, Matteo — 60
Bernardini, Chiara — 84
Berselli, Giovanni — 13
Bi, Zheyuan — 19
Bianchi, Maël — 99
Biyani, Chandra Shekhar — 64
Bonfè, Marcello — 84
Borghesan, Gianni — 84, 95
Breedon, Philip — 18
Brockdorff, Michael — 65
Brown, Alexander — 82, 97, 98
Caballero, Daniel — 58, 59
Cai, Yuyu — 17
Callera, Arianna — 33
Calmé, Benjamin — 48, 52, 65, 70, 82, 86, 97, 109
Cantillon-Murphy, Pádraig — 61
Cao, Lin — 19
Caravaca-Mora, Oscar — 72
Casals, Alicia — 57
Cattafesta, Deborah — 24
Chairez, Jorge — 112
Chandler, James H. — 48, 52, 62, 70, 76, 81, 82, 86, 97, 98, 107
Chauhan, Manish — 78, 79
Chavez-Badiola, Alejandro — 28
Chen, Hsuan-Yu — 16
Chen, Shih-Hsin — 16
Chen, Tianchi — 17
Cheng, Zhuoqi — 12, 88, 89, 103
Chikhaoui, M. Taha — 44
Chikhaoui, Taha — 27, 67
Choudhry, Omar — 64
Cianchetti, Matteo — 60
Ciuti, Gastone — 60
Cold, Kristoffer — 89
Coombes, Sean — 61
Corona Oliveira Costa , Daniel — 72
Courtecuisse, Hadrien — 99
Crowley, Daragh — 61
Cunningham John E. — 83
Damiano, Rossella — 92
Davoodi, Ayoob — 84, 95, 96
Davy, Joshua — 48, 52, 97
Dawood, Abu Bakar — 108
De Clerck, Ivo — 26
De Clerk, Ivo — 23
De Gaetano, Francesco — 33
De Mattos, Leonardo — 12
De Momi, Elena — 12, 13, 23, 55, 92
De Sanctis, Sofia — 74
de Vries, Jean-Paul P M — 41
Dean, Thomas — 52, 86
Dei, Neri Niccolo — 75
DeLorey, Charles — 65, 66
Delucchi, Tommaso — 63
Denis, Kathleen — 96
Dequidt, Jeremie — 29
Desmet, Wim — 96
Di Cosmo, Mariachiara — 24
Di Filippo, Gaia — 35
Diaz, Maria — 112
Diederichsen, Axel — 69

Diller, Eric — 41
Dimitriadis, Alexandros — 110
Dogra, Anubhav — 50
Dreyer, Juliana — 22
El Hila, Anas — 91
Erleben, Kenny — 104
Esfandiari, Mojtaba — 66
Falagario, Chiara — 12
Farola Barata, Beatriz — 72
Fasel, Lorin — 38
Fei, Jinghao — 87
Fernández-Montero, Alfonso — 58, 59
Ferraguti, Federica — 24
Finocchiaro, Martina — 104
Florent, Nageotte — 51
Fournier, Isabelle — 51
Freiling, Lukas — 40
Frieler, Simon — 41
Galindo, Gerardo — 101
Galvan, Giulia — 23, 26
Gamberini, Giulia — 53, 54
Gerig, Nicolas — 38
Getu Mhret, Teshome — 91
Giladi, Chen — 21
Gionfrida, Letizia — 11
Girerd, Cédric — 22
Gnanapragasam, Vincent — 47
Greenidge, Nikita — 48, 52, 83, 101
Guo, Xiangyu — 50
Haegemans, Joy — 26
Han, Zhe — 28
Hao, Yu — 89
Harris, Tim — 80
Harrison, Ricardo — 101
Higgins, Eoin — 61
Howell, Lewis — 62
Hu, Junlei — 50
Huber, Martin — 28
Hölker, Johannes — 39
Iacovacci, Veronica — 42, 45
Iordachita, Iulian — 66
Iovene, Elisa — 24
Ivanaj, Alice — 92
Iyama, Yuriko — 85
Jackson, Benjamin — 66
Jacobsen, Tobias — 69
Jaeger, Herman Alexander — 61
Jakobi, Jonas — 40
Jansen, Willem — 31
Jensen, Nicolai — 69
Ji, Guanglin — 104
Ji, Tianchen — 19
Jiang, Haolin — 89
Jones, Dominic — 64, 65, 66, 85
Kashef, Reza — 75
Khalid, Muhammad Abdullah — 70, 76
Kim, Do Yeon — 39, 40
Kim, Nam Gyun — 52
Kim, Sukjun — 52, 97
Konge, Lars — 89
Kumarathurai, Preman — 69
Kundrat, Dennis — 39
Leonov, Roman — 39
Li, Ruixuan — 87
Li, Xinxin — 34, 80
Liang, Ziting — 17
Liu, Junbao — 78
Liu, Xinran — 43
Lloyd, Peter — 6, 52, 81
Lodin, Carlo — 54
Lopes, Maria — 102
Lu, Chuang — 17
Lu, Haopeng — 103
Ma, Zhihao — 80
Machado, Pedro — 18
Mack, Thomas — 75
Manceau, Cécile — 22
Mandal, Nilanjan — 105
Mariano, Federico — 13
Martin Rodriguez, Ruben — 38
Marzullo, Aldo — 35, 92
Massignani, Nicola — 73
Massin, Joseph — 44
Mattos, Leonardo S. — 13
May, Yael L. — 6, 81
Mazomenos, Evangelos — 93
McLaughlan, James — 62
Menciassi, Arianna — 42, 45, 53, 54, 60
Mendizabal-Ruiz, Gerardo — 28
Mennillo, Laurent — 93
Metcalf, Adam — 65
Min Wah, Tze — 62
Misra, Sarthak — 41
Moccia, Sara — 24, 35
Mocellin, Lorenzo — 60
Mondin, Riccardo — 63
Morimoto, Tania K — 52, 97
Mosley, Connor D W — 83
Moustris, George — 110
Muradore, Riccardo — 63, 73, 74
Netz, Uri — 37
Nguyen, Bao — 86
Ning, Linxiaohai — 80
Nisky, Ilana — 37
Nixon, Josh P R — 83
Oca, Siobhan — 9
Oliveira, Bruno — 69, 89
Omran, Hassan — 32
Osman, Ali — 49
Osman, Dalia — 108
Ourak, Mouloud — 23, 26, 43, 91, 95
Ozdemir, Burak — 70, 86
Panizzolo, Riccardo — 45
Park, Shinwoo — 52

Parrotta, Francesca — 53
Pellegrini, Marco — 63
Pensabene, Virginia — 48
Peretti, Giorgio — 35
Petersen, Marianne Viola — 88
Piccinelli, Nicola — 63
Pickwell-MacPherson, Emma — 50
Pogacnik, Zeno — 72
Poulsen, Viktor — 69
Prasarnkleo, Sukrit — 95
Pérez-Salazar, Manuel J. — 58
Qi, Yuxuan — 94
Qiyu, Wei — 78
Raison, Nicholas — 11
Ramirez, Gerardo — 112
Rauter, Georg — 38
Redaelli, Alberto Cesare Luigi — 92
Regazzo, Giovanni Battista — 72
Reina, Mattia — 32
Roberti, Andrea — 63, 73, 74
Rocha, Cláudia — 31
Rocha, Manuel — 31
Rodriguez-Guerrero, Carlos — 87
Romanello, Vito — 42
Rosa, Benoit — 27, 32, 99
Ruiz Sevilla, Laura — 35
Ryu, Jee Hwan — 52
Rüegg, Calvin — 38
S. De Mattos, Leonardo — 35
Sadati, S.M.Hadi — 75
Sampieri, Claudio — 35
Sarikaya, Duygu — 49
Savarimuthu, Thiusius R. — 69, 88, 103
Sayols, Narcís — 57
Schneider, Jurgen E. — 6, 81
Shabanian, Daniel — 18
Sheikh Sofla, Mohammad — 75
Shilon, Lior — 21
Siepel, Françoise — 31
Silvestri, Giuseppe — 28
Sousa, Armando — 31
Srymbetov, Tamerlan — 42, 45
Stabile, Giovanni — 53
Stefanini, Cesare — 60
Stilli, Agostino — 17
Stoyanov, Danail — 17
Sun, Haorui — 79
Suulker, Cem — 75, 108
Sánchez-Margallo, Francisco M. — 56, 58, 59
Sánchez-Margallo, Juan A. — 56, 58, 59
Thieringer, Florian — 38
Thompson, Alexander J — 105
Tian, Chuan — 89
Tian, Huanyu — 28
Tian, Sizhe — 29
Timmermans, Maikel — 96
Tognarelli, Selene — 53, 54
Toilos, Christos — 93
Tombini, Ingrid — 55
Tse Zion — 34, 80
Tse, Zion — 47, 94
Tzafestas, Costas — 110
Ulrich, Ambroise — 67
Valdastri, Pietro — 6, 48, 50, 52, 65, 66, 70, 76, 81, 82, 83, 85, 86, 97, 98, 101, 107, 109
van Alphen, Maarten — 31
Van Assche, Kaat — 84, 96
Van den Bogaert, Marthe — 96
van den Brekel, Michiel — 31
Vander Poorten, Emmanuel — 23, 26, 43, 72, 84, 87, 91, 95, 96, 102
Venkiteswaran, Venkatasubramanian Kalpathy — 41
Vercauteren, Tom — 28
Vilaseca, Isabel — 35
Wachtel, Lukas — 105
Walia, Kartikeya — 18
Waterston, Manisha — 62
Wei, Qiyu — 79
Williams, Jackson — 105, 107
Wolff, Jan E H — 39
Wright, Sasha-Gay — 101
Wu, Di — 103
Wu, Diyang — 19
Xie, Xiao — 51
Yanez Trujillo, Andrea Libertad — 48, 105
Yeh, Keng-Hua — 16
Yin, Hang — 104
Yvert, Blaise — 67
Zemiti, Nabil — 22
Zeng, Lingyun — 28
Zhang, Yao — 91
Zhang, Zhenyu — 108
Zheng, Gang — 51
Öztürk, Cengizhan — 46, 68
Şahin, İsmail — 46, 68

CRAS 2026



15th Conference on New Technologies for Computer and Robot-Assisted Surgery

16–18 SEPTEMBER 2026

University of Leeds · United Kingdom

UNIVERSITY OF LEEDS

ROBOTICS AT LEEDS

cras-eu.org

LEEDS · UK