Control of Multi-Degree-of-Freedom Catheters in Unknown Environments
Exploring the Potential of Deep Learning and Augmented Reality

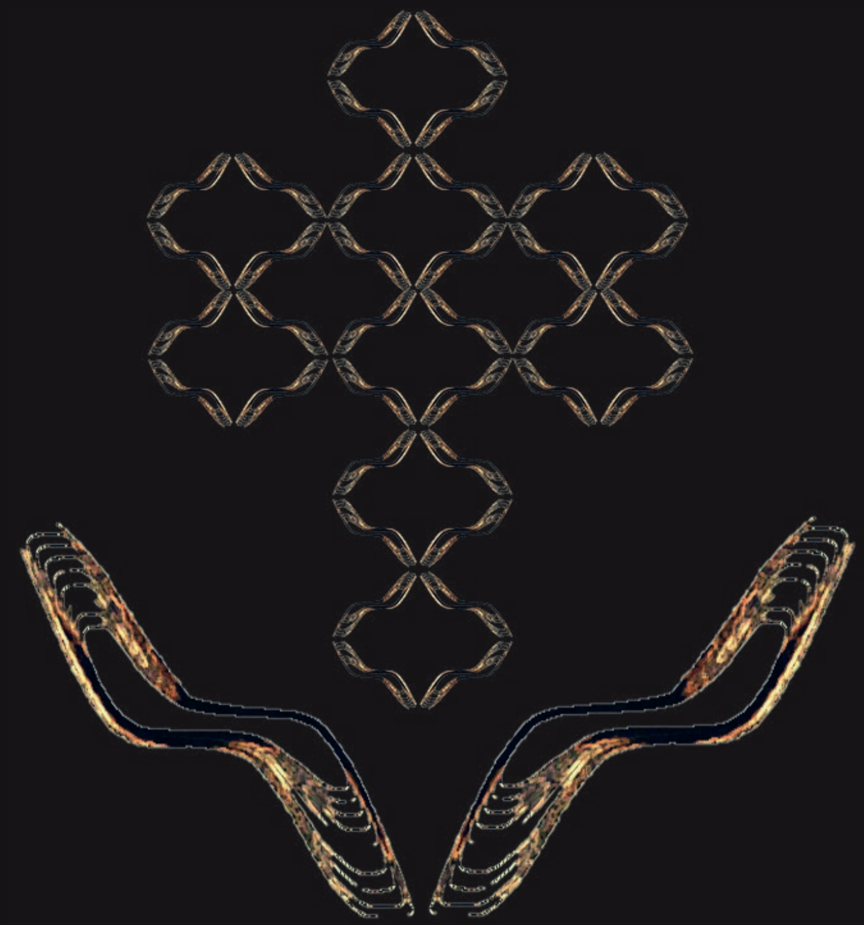
Wu, D.

DOI
10.4233/uuid:21af38b1-ae31-4247-adf2-dcbff8b9ebb1

Publication date
2023

Document Version
Final published version

Citation (APA)
Wu, D. (2023). *Control of Multi-Degree-of-Freedom Catheters in Unknown Environments: Exploring the Potential of Deep Learning and Augmented Reality*. [Dissertation (TU Delft), Delft University of Technology, Katholieke Universiteit Leuven]. https://doi.org/10.4233/uuid:21af38b1-ae31-4247-adf2-dcbff8b9ebb1



TU Delft Delft University of Technology

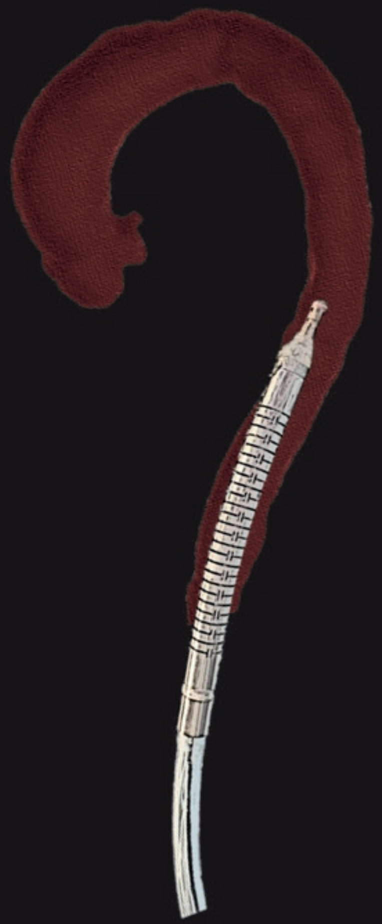
KATHOLIEKE UNIVERSITEIT LEUVEN

Control of Multi-DOF Catheters in Unknown Environments

Di Wu

Control of Multi-Degree-of-Freedom Catheters in Unknown Environments

Exploring the Potential of Deep Learning and Augmented Reality



Di Wu

Control of multi-degree-of-freedom catheters in unknown environments

EXPLORING THE POTENTIAL OF DEEP LEARNING AND AUGMENTED REALITY

Control of multi-degree-of-freedom catheters in unknown environments

EXPLORING THE POTENTIAL OF DEEP LEARNING AND AUGMENTED REALITY

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Wednesday 13 December 2023 at 17:30 o'clock

by

Di WU

Master of Science in Medical Technology and Engineering,
Technische Universität München, Germany,
born in Anhui, China.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. J. Dankelman,	Delft University of Technology, promotor
Dr. ir. E. Vander Poorten,	Katholieke Universiteit Leuven, promotor
Independent members:	
Prof. dr. ir. P. Breedveld,	Delft University of Technology
Prof. dr. ir. J. Vander Sloten,	Katholieke Universiteit Leuven, Belgium
Prof. dr. ir. P. Cantillon-Murphy,	University College Cork, Ireland
Dr. ir. A. Sakes,	Delft University of Technology
Prof. dr. ir. J. van den Dobbelsteen,	Delft University of Technology, reserve member
Other member:	
Dr. ir. M. Ourak,	Katholieke Universiteit Leuven, Belgium

The doctoral research has been carried out in the context of an agreement for double degree PhD program between Delft University of Technology, the Netherlands and Katholieke Universiteit Leuven, Belgium.

The research presented in this dissertation was supported by the ATLAS project, which received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813782. This research was also funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017140 for the ARTERY project. Further support was provided by the FLEXFET project (3E200384), an internally funded C2 project at KU Leuven.

Keywords: robotic catheter, deep learning, control, sensing, augmented reality

Printed by: Gildeprint

ISBN 978-94-6384-521-2

To my beloved family and cherished ones
You are the essence of my joy and the sanctuary of my heart

Abstract

Cardiovascular disease is currently one of the biggest threats to health. Specific types of cardiovascular disease include, but are not limited to, coronary artery disease, cardiac valve disorders, or peripheral arterial disease. The current gold standard for managing these conditions incorporates the use of catheters and guidewires for intravascular navigation. Following their insertion into the vascular system, these instruments facilitate a variety of procedures, such as stent placement, recanalization of vessel blockage, and radiofrequency ablation. Compared to more invasive open heart surgery, catheterization represents a minimally invasive approach. This offers several benefits, including smaller incisions, faster postoperative recovery, and improved aesthetic outcomes.

However, current procedures still present considerable challenges. First, the environment through which the instruments are to be navigated is highly complex. The structure of blood vessels is not only intricate and delicate, but also prone to deformation. Catheterization is further complicated by the dynamic nature of the environment, influenced by physiological movements such as heartbeat, respiration, and blood flow. Second, the existing procedure relies on fluoroscopy for visual feedback, which offers only two-dimensional grayscale images, thus physicians find it challenging to assess depth information from such an image. More importantly, it is an ionizing radiative imaging modality that introduces hazards to both patients and clinicians. The risks are especially high for clinicians, given their repeated exposure across procedures. Furthermore, inherent nonlinearity and hysteresis due to material properties or friction inside the actuators of catheter as well as the continuum/soft structure of the catheter cannot be neglected. Hysteresis complicates precise catheter control, especially during reversal of motion. Energy may build up when the catheter is stuck. Upon release, this could potentially result in excessive forces applied by the acute catheter tip to the vessel and could potentially lead to vascular perforation. Lastly, the entire procedure is still performed manually by clinicians, making it highly skill-intensive and characterized by a long learning curve. Managing all complexities simultaneously is very mentally demanding for a human operator. Robotic assistance techniques could potentially assist in parts of the procedure and as such improve these interventional approaches.

To address the aforementioned issues, the thesis develops dedicated types of robotic assistance to enhance the safety and precision of catheterization procedures. A key focus is on the modeling and control of robotic catheter. In addition, recent technologies such as Deep Learning (DL) and Augmented Reality (AR) are incorporated. DL, given its ability to model highly non-linear relations, is applied in the modeling and control of robotic catheters. Augmented reality, as a three-dimensional display technology, is used to enhance visual guidance for physicians.

First, to provide some background, a comprehensive survey of machine learning (ML) applications is conducted across all Flexible Surgical and Interventional Robots (FSIR). In this survey, a 2D classification method is proposed, intended to categorize the wide array of Machine Learning (ML) algorithms, which encompasses both traditional ML methods and the latest DL approaches that have been utilized so far in the realm of FSIR. Subsequently, the applications of ML algorithms are discussed per specific subdomains: perception, modeling, control, and navigation. The trends, popularity, strengths, and limitations of various ML algorithms are compared and analyzed. Finally, discussions are held on the current limitations and future prospects of ML techniques in robotically-assisted Minimally Invasive Procedures (MIP).

In order to overcome the limitations of fluoroscopy, efforts are made to incorporate teleoperation and enhanced visualization. To achieve this, a range of devices can be utilized. In this work, gaming joysticks and AR Head-Mounted Displays (HMD) have been used to realize teleoperation, while AR HMD and a 2D standard monitor are employed to enhance visualization. Various interactive modes for intuitive human-in-the-loop catheter steering have been proposed. The effectiveness of different interactive modes is investigated through an *in-vitro* user study. User performance is evaluated both subjectively and objectively.

Third, to ease the clinician's awareness of the interventional scene, additional work is done to capture the catheter's shape in a non-radiative fashion. Hereto, a DL method for shape sensing of continuum robots based on multicore FBG fiber is introduced. The proposed method allows shape sensing based on a multicore fiber that is placed off-center. By avoiding the need to place shape sensors at the center of the catheter, it becomes possible to miniaturize the catheter and leave the central channel available for other purposes. A 2-DOF bench-top fluidics-driven catheter system is built to validate the proposed ANN. The proposed approach is compared with a recent state-of-the-art model-based shape-sensing approach. Exhaustive ablation studies have been conducted to enhance the understanding of ANN, thereby increasing its interpretability. The proposed method is shown to be able to precisely reconstruct the shape of a robotic catheter without needing to resort to radiation-based techniques such as fluoroscopy. The precise shape reconstruction can offer valuable information that could also be beneficial in controlling the catheter.

Fourth, in addition to improving the knowledge about where the catheter is, contributions have been made to improve the precision of catheter positioning. A DL-based method is proposed to model the hysteresis of a robotic catheter. To provide a more comprehensive assessment of the DL method, an analytical model and a conventional ML method have been implemented for comparison. Subsequently, the proposed method was also validated through reconstruction of an anatomy using a forward-looking imaging modality. The proposed method was found to accurately predict the response of the catheter tip solely based on the proximal input pressure, despite the presence of large amounts of hysteresis. This knowledge could help avoid the need to install sensors at the tip of the catheter for localization, which is challenging due to spatial limitations and sterilization requirements.

Fifth, a DL-based position controller for the robotic catheter is introduced. The controller compensates for hysteresis and captures the inverse kinematics of the catheters, thereby paving the way for precise position control. Building on this position controller, a novel compliant motion control algorithm was introduced. The compliant controller automatically minimizes the interaction force between the catheter tip and its surroundings, thereby facilitating safer interventions. This approach also underscores the potential of leveraging DL to avoid exhaustive modeling traditionally needed for accurately navigating continuum robots in intricate environments such as the patient's vasculature.

The primary objective of this thesis is to enhance the precision and intuitive control of robotic catheters for safe navigation in complex environments. This work predominantly utilizes Deep Learning (DL) and Augmented Reality (AR), applying these technologies across various research objectives. The following contributions have been achieved: 1. conducting a comprehensive survey to provide a solid research background for this thesis. 2. employing AR technology to improve visualization and control during human-in-the-loop teleoperated catheterization. Different interactive modes are implemented by integrating an AR Head-Mounted Display (HMD), a gamepad, or a standard monitor, and these modes are compared. 3. enhancing clinicians' awareness of the catheter 3D shape through a non-radiative method and improving the accuracy of shape reconstruction using DL techniques. 4. utilizing DL to model and compensate for the hysteresis present in robotic catheters. 5. elevating the level of autonomy in catheter control, with a demonstration in an *in-vitro* environment.

Beknopte samenvatting

Cardiovasculaire ziekten vormen momenteel een van de grootste bedreigingen voor de volksgezondheid. Specifieke typen omvatten onder andere coronaire hartziekten, hartklepaandoeningen en perifere arteriële ziekten. De huidige gouden standaard voor het behandelen van deze aandoeningen omvat het gebruik van katheters en geleidedraden voor intravasculaire navigatie. Deze instrumenten vergemakkelijken na inbrenging in het vaatstelsel diverse procedures, zoals stentplaatsing, hercanalisatie van vaatblokkades en radiofrequente ablatie. Katheterisatie is een minimaal invasieve benadering en biedt diverse voordelen, zoals kleinere incisies, sneller postoperatief herstel en betere esthetische resultaten.

Er zijn echter nog aanzienlijke uitdagingen. Ten eerste is de omgeving waarin de instrumenten moeten worden genavigeerd uiterst complex. De structuur van bloedvaten is niet alleen ingewikkeld en delicaat, maar ook onderhevig aan vervorming. Daarnaast wordt katheterisatie bemoeilijkt door de dynamische aard van de omgeving, beïnvloed door fysiologische bewegingen zoals hartslag, ademhaling en bloedstroom. Ten tweede is voor visuele feedback fluoroscopie nodig, wat alleen tweedimensionale grijsschaalbeelden biedt, waardoor het moeilijk is voor artsen om diepte-informatie te beoordelen. Belangrijker nog, het is een ioniserende beeldvormingsmodaliteit die gevaren oplevert voor zowel patiënten als clinici. Vooral voor clinici zijn de risico's hoog, gezien hun herhaalde blootstelling aan procedures.

Om de genoemde problemen aan te pakken, ontwikkelt het proefschrift specifieke soorten robotische assistentie om de veiligheid en precisie van katheterisatie-procedures te vergroten. Er is een belangrijke focus op de modellering en besturing van de robotische katheter. Daarnaast worden recente technologieën zoals Deep Learning (DL) en Augmented Reality (AR) geïntegreerd. DL, vanwege het vermogen om zeer niet-lineaire relaties te modelleren, wordt toegepast bij de modellering en besturing van robotische katheters. Augmented reality, als driedimensionale weergavetechnologie, wordt gebruikt om visuele begeleiding voor artsen te verbeteren.

Eerst wordt er, om wat achtergrond te bieden, een uitgebreid overzicht uitgevoerd van machine learning (ML) toepassingen over alle Flexibele Chirurgische en

Interventierobots (FSIR). In dit overzicht wordt een 2D-classificatiemethode voorgesteld om het brede scala aan machine learning (ML) algoritmen te categoriseren, waaronder zowel traditionele ML-methoden als de nieuwste DL-benaderingen die tot nu toe in het domein van FSIR zijn gebruikt. Vervolgens worden de toepassingen van ML-algoritmen per specifieke subdomeinen besproken: perceptie, modellering, besturing en navigatie. De trends, populariteit, sterke en zwakke punten van verschillende ML-algoritmen worden vergeleken en geanalyseerd. Ten slotte worden discussies gevoerd over de huidige beperkingen en toekomstige vooruitzichten van ML-technieken bij robotgestuurde minimaal invasieve procedures (MIP).

Ten tweede worden, om de beperkingen van fluoroscopie te overwinnen, inspanningen geleverd om teleoperatie en verbeterde visualisatie te integreren. Hiertoe kunnen verschillende apparaten worden gebruikt. In dit werk zijn game-joysticks en AR Head-Mounted Displays (HMD) gebruikt om teleoperatie te realiseren, terwijl AR HMD en een 2D-standaardmonitor worden gebruikt om de visualisatie te verbeteren. Verschillende interactieve modi voor intuïtieve besturing van de katheter door de mens zijn voorgesteld. De effectiviteit van de verschillende interactieve modi wordt onderzocht door middel van een in-vitro gebruikersstudie. Gebruikersprestaties worden zowel subjectief als objectief geëvalueerd.

Ten derde wordt extra werk verricht om het bewustzijn van de clinicus over de interventieomgeving te vergroten. Hiertoe wordt een DL-methode voor vormsensing van continuümrobots geïntroduceerd op basis van multicore FBG-vezel. De voorgestelde methode maakt vormdetectie mogelijk op basis van een enkele multicore-vezel die excentrisch is geplaatst. Door de noodzaak te vermijden om vormsensoren in het midden van de katheter te plaatsen, wordt het mogelijk om de katheter te verkleinen en het centrale kanaal vrij te laten voor andere doeleinden. Een 2-DOF benchtop vloeistofaangedreven kathetersysteem is gebouwd om de voorgestelde ANN te valideren. De voorgestelde aanpak wordt vergeleken met een recente state-of-the-art modelgebaseerde vormdetectiemethode. Er zijn uitgebreide ablatiestudies uitgevoerd om het begrip van ANN te vergroten, waardoor de interpreteerbaarheid ervan toeneemt. De voorgestelde methode blijkt in staat om de vorm van een robotische katheter nauwkeurig te reconstrueren zonder gebruik te hoeven maken van beeldvormingstechnieken gebaseerd op straling zoals fluoroscopie. De precieze vormreconstructie kan waardevolle informatie bieden die ook gunstig kan zijn om de katheter te besturen.

Ten vierde zijn, naast verbeterde kennis over de locatie van de katheter, bijdragen geleverd om de precisie van de katheterpositionering te verbeteren. Er wordt een DL-gebaseerde methode voorgesteld om de hysterese van een robotische katheter te modelleren. Om een meer uitgebreide beoordeling van de DL-methode te bieden, zijn een analytisch model en een conventionele ML-methode geïmplementeerd voor vergelijking. Vervolgens wordt de voorgestelde methode ook gevalideerd door de reconstructie van een anatomie met behulp van een voorwaarts kijkende beeldvormingsmodaliteit. Er wordt vastgesteld dat de voorgestelde methode nauwkeurig de reactie van de kathetertip kan voorspellen op basis van enkel proximale invoerdruk, ondanks de aanwezigheid van grote hoeveelheden hysterese. Deze kennis kan helpen om de noodzaak te vermijden om sensoren op de punt van de katheter te installeren voor lokalisatie, wat uitdagend is vanwege de ruimtelijke beperkingen en sterilisatie-eisen.

Ten vijfde wordt een op deep learning gebaseerde positiecontroller van de robotische katheter geïntroduceerd. De controller compenseert voor hysterese en legt de inverse kinematica van de katheters vast, waardoor de weg wordt vrijgemaakt voor precieze positiebesturing. Voortbouwend op deze positiecontroller wordt een nieuwe conforme bewegingsregelingsalgoritme geïntroduceerd. De conforme controller minimaliseert automatisch de interactiekracht tussen de punt van de katheter en de omgeving, waardoor veiligere interventies mogelijk zijn. Deze aanpak benadrukt ook het potentieel van het gebruik van DL om uitputtende modellering te vermijden die traditioneel nodig is voor het nauwkeurig navigeren van continuümrobots in complexe omgevingen zoals het vaatstelsel van de patiënt.

Het primaire doel van deze scriptie is het verbeteren van de precisie en intuïtieve controle van robotische katheters voor veilige navigatie in complexe omgevingen. Dit werk maakt voornamelijk gebruik van Deep Learning (DL) en Augmented Reality (AR), waarbij deze technologieën worden toegepast op verschillende onderzoeksdoelstellingen. De volgende bijdragen zijn bereikt: 1. het uitvoeren van een uitgebreide enquête om een solide onderzoeksachtergrond voor deze scriptie te bieden. 2. het toepassen van AR-technologie om de visualisatie en controle tijdens door mensen bediende telegeleide katheterisatie te verbeteren. Verschillende interactieve modi worden geïmplementeerd door het integreren van een AR Head-Mounted Display (HMD), een gamepad of een standaardmonitor, en deze modi worden vergeleken. 3. het verhogen van het bewustzijn van clinici over de 3D-vorm van de katheter door middel van een niet-radiatieve methode en het verbeteren van de nauwkeurigheid van vormreconstructie met behulp van DL-technieken. 4. het gebruik van DL om het hystereseverschijnsel dat aanwezig is in robotische katheters te modelleren en te compenseren. 5. het verhogen van het niveau van autonomie in kathetercontrole, met een demonstratie in een in-vitro-omgeving.

List of Abbreviations

2D	2 dimensional
3D	3 dimensional
ANN	artificial neural networks
AR	augmented reality
CNN	convolutional neural network
CPU	central processing unit
CT	computed tomography
CTR	concentric tube robot
CVD	cardiovascular disease
DL	deep learning
DoF	degree(s) of freedom
DRL	deep reinforcemnt learning
EDM	electrical discharge machining
EIT	electrical impedance tomography
EM	electromagnetic
EVAR	endovascular aneurysm repair
F/T	force/torque
FBG	fiber bragg grating
FK	forward kinematics
FNN	feedforward neural network
FSIRs	flexible surgical and interventional robots
GMM	Gaussian mixture model
GMR	Gaussian mixture regression
GPI	generalized Prandtl-Ishlinskii
HMD	head-mounted display

IDRDPI	inverse deadband rate-dependent Prandtl-Ishlinskii
IK	inverse kinematics
IVUS	intravascular ultrasound
MADQN	multiagent deep Q network
MAE	maximum absolute error
MIPs	minimally invasive procedures
ML	machine learning
MoAE	mean of the absolute error
MRI	magnetic resonance imaging
NRMSE	normalized root mean square error
OCT	optical coherence tomography
PAM	pneumatic artificial muscle
PID	proportional–integral–derivative
RAM	random-access memory
ReLU	rectified linear unit
RL	reinforcement learning
RMSE	root mean square error
RNN	recurrent neural network
ROS	robot operating system
SGC	steerable guide catheter
SVD	singular value decomposition
SVM	support vector machine
VR	virtual reality
WHO	world health organization

List of Symbols

Chapter 3

${}^{E}\boldsymbol{p}$	positions of the ten markers represented in EM frame
${}^{M}\boldsymbol{p}$	positions of the ten markers represented in mesh model frame
$\boldsymbol{R}$	optimal rotation matrix obtained through SVD method
$\boldsymbol{s_i^d}$	desired trajectories
$\boldsymbol{s_n^d}$	position of the target
$\boldsymbol{s_j}$	actual trajectories
${}^{E}\boldsymbol{T_C}$	transformation matrix from catheter tip frame to EM frame
${}^{M}\boldsymbol{T_C}$	transformation matrix from catheter tip frame to mesh model frame
${}^{M}\boldsymbol{T_E}$	transformation matrix from EM frame to mesh model frame
Δ	direction switching button on the gamepad
ϕ	bending angle of the left thumbstick in y axis
θ	bending angle of the left thumbstick in x axis
$\{\boldsymbol{C}\}$	frame representation of the catheter tip
$\{\boldsymbol{E}\}$	frame representation of the EM sensor
$\{\boldsymbol{M}\}$	frame representation of the mesh model
p_ϕ	pressure applies on catheter y axis
p_θ	pressure applies on catheter x axis
r	bending angle of the right thumbstick
T	duration of the experiment

t_0	starting time
T_a	targeting error
t_g	the time when stopping the timer
T_r	tracking error
v	velocity of the catheter driver

Chapter 4

$\Delta\lambda$	wavelength shift
κ	catheter curvature
κ_f	calculated curvature
κ_{gt_i}	ground truth curvature that is generated from EM poses
λ	measured wavelength
λ_0	unstrained Bragg wavelength
$\overrightarrow{OK}$	bending direction vector
$\overrightarrow{t_{Base}}$	base tangent vector
$\overrightarrow{t_{Tip}}$	tangent vector of the tip EM sensor
τ	torsion on the catheter
θ_f	angle between $\overrightarrow{OO_f}$ and the x-axis
C	Cartesian position of each point that belongs to the catheter length
c	distance between the tip and the base position that equals to the chord length
d_f	distance to the center line
l	length of the bendable segment
O	center of the circle that fits the catheter's shape
$O_f^{'}$	projection of O_f on the bending direction vector $\overrightarrow{OK}$
O_f	center of the fiber
P^{GT}	heat shrinks' locations in the ground truth frame
P^{Img}	heat shrinks' locations in the image frame
P_{Base}	catheter base position

P_{Tip}	catheter tip position
r	radius of the circle that fits the catheter's shape

Chapter 5

$*$	element-wise multiplication
ϵ	margin of tolerance
γ	kernel coefficient
σ	sigmoid function
τ	factor that regulates descending speed of the sine wave
θ	bending angle of the catheter tip
$\widetilde{c}^{(t)}$	a candidate value for the cell state update
A	amplitude of the sine wave
b_f	bias for the forget gate
b_g	bias for creating the new cell state
b_i	bias for the input gate
b_o	bias for the output gate
C	panelty
c	chirpyness
$c^{(t)}$	updated cell state
$c^{(t-1)}$	previous cell state
$d^{'}$	displacement of the catheter tip
f	frequency of the sine wave
$f^{(t)}$	output of the forget gate
H	distance from the laser sensor to the catheter base along the direction when the catheter is straight
$h^{(t)}$	final output of the LSTM cell for the current time step
$h^{(t-1)}$	output of the LSTM cell from the previous time step
H_0	null hypothesis
$i^{(t)}$	Output of the input gate

$i^{(t)}$	output of the input gate
k	number of folds in cross-validation
L	entire length of the catheter
m	slope of the relation between the s and U
n	y-intercept of the relation between the s and U
$o^{(t)}$	output of the output gate
$p^{(t)}$	current input pressure
p_b	vector representing the base direction derived from FBG
p_t	vector representing the tip direction derived from FBG
p_1	descending sinusoidal pressure with zero baseline
p_2	descending sinusoidal pressure without zero baseline
s	distance from the laser sensor to the object
U	output voltage of the laser sensor
U_f	recurrent weight matrix associated with the forget gate
U_g	recurrent weight matrix associated with the new cell state
U_i	recurrent weight matrix associated with the input gate
U_o	recurrent weight matrix associated with the output gate
V	voltage sent to pressure valve
W_f	weight matrix associated with the forget gate
W_g	weight matrix associated with the new cell state
W_i	weight matrix associated with the input gate
W_o	weight matrix associated with the output gate

Chapter 6

A	amplitude of the sine wave
τ	factor that regulates descending speed of the sine wave
θ	bending angle of the catheter
B	offset amplitude

c	chirpyness
f	frequency of the sine wave
K_{Off}	catheter stiffness when the compliant motion controller is off
K_{on}	catheter stiffness when the compliant motion controller is on

Contents

Chapter 1

Introduction

1.1 Evolution of surgery

Surgical procedures have played a crucial role in saving people. It has revolutionized the field of healthcare, improving the quality of life of many. Traditional "open" surgical approaches have been increasingly replaced by minimally invasive procedures since the mid-20th century. Examples include endoscopic surgery, which originated in the 19th century [1], and laparoscopic surgery, which gained prominence in the late 1980s [2]. Laparoscopic surgery, which is an example of the so-called "keyhole surgery", is a revolutionary minimally invasive procedure that allows surgeons to reach the abdomen or pelvis without large incisions. Using a series of small holes, typically not exceeding a centimeter in diameter, surgeons introduce a laparoscope, a slender tube fitted with a high-resolution camera, and a high-intensity illuminator, for visualization of the surgical area [3, 4]. Dedicated instruments are subsequently manipulated through these keyholes to execute the procedure. The minimally invasive nature of this method contributes to faster patient recovery, reduces postoperative discomfort, and minimizes scarring, explaining its popularity in a wide range of surgeries.

Although minimally invasive surgery marked a significant step forward, it came with its own set of difficulties. Surgeons have a limited view and often operate in a non-ergonomic posture, making it a physically demanding procedure, leading to both mental and physical fatigue. Furthermore, natural hand tremors tend to be amplified and, as such, could affect the accuracy of operations. These obstacles led to the development of robotic-assisted surgery. The Da Vinci Surgical System from Intuitive Surgical Inc., introduced in the early 2000s, represents a significant milestone in this field. The Da Vinci robot offers several notable advantages, including enhanced precision, superior visual feedback, and a reduction in the physical workload of surgeons [5]. Furthermore, it popularized a novel paradigm, teleoperation-based surgery, paving the way for other players in this field.

Note that robotic laparoscopic surgery still requires making 2-4 incisions on the patient's abdomen. Therefore, small or even non-invasive approaches are being investigated to further reduce damage, such as endoluminal procedures or vascular interventions. The challenge of these procedures lies in skillfully maneuvering a flexible instrument by clinicians through natural orifices (e.g., mouth, urethra, anus, vagina, nose) or a small incision to reach the target area. Robotic assistance is also increasingly being applied in this field, exemplified by systems such as the Corindus vascular robot (Siemens Healthineers, Germany). The control of the Corindus system is achieved via teleoperation. The physician operates from a control room while a catheter driver is stationed next to the patient to manipulate the catheters. This approach reduces radiation exposure for the physician.

Despite the progress made in the field of surgical and interventional robots, overall progress has been slow. Factors such as the complexity of the environment, the challenge of accurately managing the force applied to delicate tissues, and the high associated risks contribute to this slower development. Therefore, in recent years, people have started to apply Machine Learning (ML) techniques to surgery. Thanks to advances in computing power and algorithm development, ML and its subset, Deep Learning (DL), are widely used for surgical image recognition and surgical data analysis. Augmented Reality (AR) technology is also being applied to surgery, providing surgeons with enhanced visual feedback. It is imperative to continue to innovate in this field, especially considering that advancement can directly contribute to human health.

1.2 Clinical background

1.2.1 Cardiovascular disease

Cardiovascular disease (CVD) is a common health problem. CVD forms the leading cause of death worldwide. According to the World Health Organization (WHO) statistics, approximately 17.9 million people died from CVDs in 2019, making up 32% of all deaths worldwide. These numbers are expected to increase to 23.6 million by 2030 [6]. CVD results from a combination of factors that include unhealthy lifestyle habits such as poor diet, sedentary behavior, smoking, and excessive alcohol consumption, along with health conditions such as high blood pressure, diabetes, and high cholesterol [7, 8]. CVDs encompass a diverse range of conditions affecting the heart and blood vessels. This broad category of CVDs includes coronary artery disease, stroke, irregular heart rhythms, aortic aneurysms, valve disease, and peripheral artery disease [8, 9]. Each of these conditions poses its own unique challenges and impacts the overall health of the heart and circulatory system.

When left untreated, CVDs can lead to severe complications. These include heart failure, heart attacks, and strokes, resulting from impaired pumping capacity of the heart or from blood clots obstructing circulation. Aneurysms pose a risk of life-threatening internal hemorrhage [10], while peripheral artery disease can manifest as pain in the limbs [11]. Cognitive problems such as vascular dementia, caused by impaired blood flow to the brain, can also occur [12]. Furthermore, it can significantly affect the daily lives of patients, leading to physical limitations, mental health issues such as anxiety and depression, substantial financial burden from medical costs, and overall decrease in quality of life [13]. Above alarming statistics combined with profound effects on patient quality of life, underscore the urgent need for effective treatments and innovative strategies in the management of CVD.

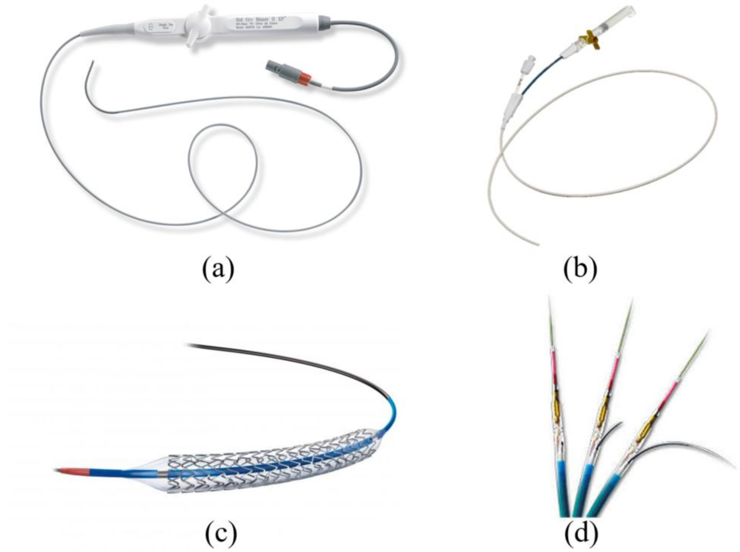


Figure 1.1: Different types of commercially available catheters: (a) Blazer™ Open-Irrigation Ablation Catheter, Boston scientific Corporation; (b) Occluder™ Occlusion Balloon Catheter, Boston Scientific Corporation; (c) Stent catheter, Boston Scientific Corportation; (d) Pioneer Plus™ IVUS guided re-entry catheter, Phillips Inc.

1.2.2 Catheterization

Catheterization, specifically endovascular catheterization, is a clinical procedure used to diagnose and treat heart conditions. Over 3.5 million procedures are conducted in Europe each year, with the number continuing to rise [14]. Catheterization involves the insertion of a flexible slender tube, known as a catheter, through a blood vessel. Figure 1.1 displays various types of commercially available catheters, each with different functionalities, used in different procedures. The catheterization procedure allows physicians to perform diagnostic tests and treatments without having to perform open chest surgery. The procedure is usually performed in a hospital catheterization laboratory, abbreviated as “cath lab” as shown in Fig. 1.2. The patient is given a local anesthetic. A small incision is made, usually in the groin, arm, or neck, where a blood vessel is close to the surface of the skin [15]. In most cases, a guidewire is first inserted into this incision and then meticulously navigated until it reaches the target area. The guidewire acts as a pathway to guide other instruments, e.g., catheters, safely through the blood vessels. The position and motion of the guidewire/catheter can be tracked using fluoroscopy, which is acquired through a C-arm system. This setup (see Fig. 1.2) offers real-time feedback, crucial for ensuring precise and accurate navigation during procedures.

Once the catheter has been inserted and placed, it opens opportunities for a variety of procedures. One such procedure is angioplasty that involves the injection of a radiopaque dye through the catheter. This is a key imaging technique that provides a detailed visualization of blood vessels, which can be instrumental in diagnosing cardiovascular conditions such as coronary artery disease. The internal channel of the catheter can be utilized to introduce instruments that are designed for therapeutic purposes. For example, in angioplasty, a tiny balloon is introduced and then expanded to widen the arteries that have become constricted or blocked, thus improving blood circulation. In addition, stents, which are small mesh-like tubes, can be inserted through the catheter to prevent the arteries from narrowing again, thus ensuring stable and healthy blood flow.

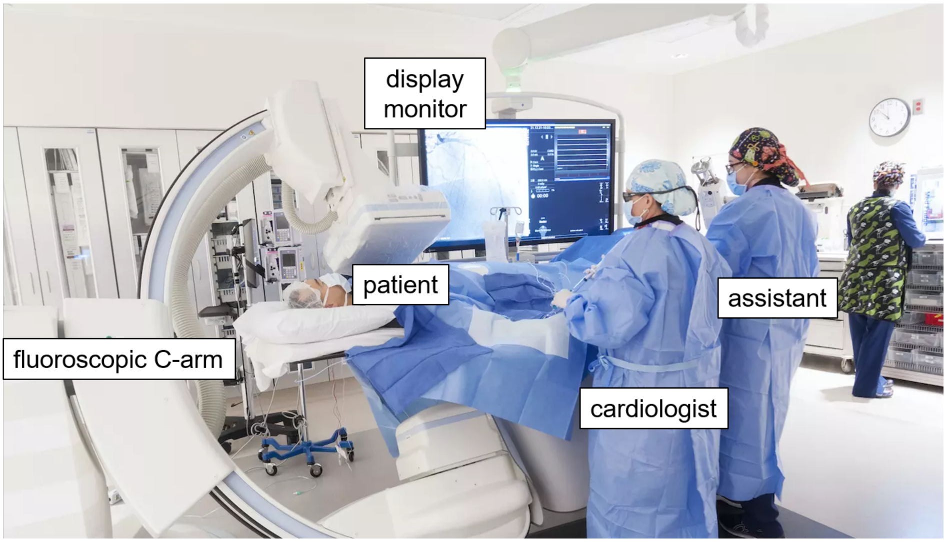


Figure 1.2: The current practice for endovascular interventions in a catheterization laboratory ("cath lab") typically involves a cardiologist stand next to the patient. In this setup, the cardiologist steers the catheter or guidewire, guided by fluoroscopic imaging obtained through a C-arm [16].

In addition to these, the catheter can also be used to perform intricate procedures such as cardiac ablation. This treatment specifically addresses abnormal heart rhythms by directing energy through a manually-steerable catheter. The resulting scars on the heart tissue function as barriers, effectively interrupting the irregular electrical signals that cause abnormal rhythm. Cardiac ablation can restore the heart rhythm to a normal state.

Furthermore, the catheter is a vital component in the execution of treatments designed to repair malfunctioning heart valves, for example, the MitraClip approach

[17, 18]. This procedure is commonly used in patients with mitral regurgitation, a condition characterized by the inability of the mitral valve of the heart to seal tightly. Due to this malfunction, blood flows back into the heart. Therefore, the heart will need to work excessively hard to ensure that enough blood enters the circulation system. During the MitraClip procedure, a catheter is used to attach a clip to the compromised mitral valve, preventing the valve from opening during ventricular systole. This intervention helps reduce atrial regurgitation, thus improving the efficiency of heart pumping function.

1.3 Procedure challenges and limitations

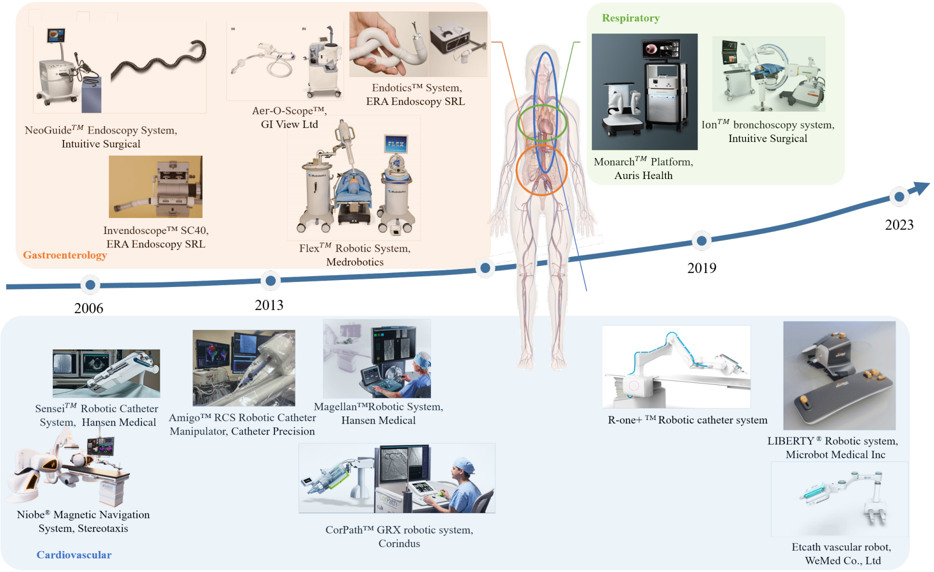


Figure 1.3: Endoluminal robotic systems that are commercially available or have received approval from relevant authorities for market release. These systems are categorized into three specialties, i.e., gastroenterology, respiratory, and cardiovascular.

Endovascular interventions face two major challenges: *controllability* and *awareness* of the instrument. Controllability focuses on the physical manipulation of catheters and guidewires, while awareness refers to the use and interpretation of imaging guidance.

Challenges in controllability of instrument: The intricate and fragile nature of the human vasculature makes the navigation of guidewires, catheters, or guidance sheaths complex and challenging. Inherent factors in the catheter itself, such as hysteresis, buckling, slack, and plastic deformation, further complicate precise control. Current procedures are mostly carried out manually [19], which requires high levels of skill and experience. Manual control could result in a lack of consistency and reliability, as it is heavily based on the expertise of the individual physician.

Additionally, the clinicians performing the procedure are exposed to ionizing radiation, which motivates the development of remote or teleoperated approaches. However, teleoperation has its own challenges. For example, it can become more difficult to maintain correct hand-eye coordination when the operator is not physically present at the operational site. Furthermore, the absence of real-time haptic (touch) feedback makes it challenging to gauge the interaction between the catheter and the patient's anatomy, increasing the risk of unintentional damage or complications [20, 21]. Reliance then shifts heavily towards visual cues, which do not provide an easy understanding of the governing interaction forces [22].

Challenges in awareness of instrument: Fluoroscopy, the imaging modality most often employed during endovascular interventions, produces 2D grayscale images. This imaging modality limits the quality of visual information, as it is difficult to perceive depth and spatial relationships between structures. Physicians need an in-depth understanding of the anatomy of the cardiovascular patient to accurately correlate catheter position and movements with its surrounding environment. The radiative nature of fluoroscopy also raises concerns about repeated or prolonged exposure to ionizing radiation, both for the patient and the healthcare professionals involved [23, 24]. Although physicians can wear lead aprons to mitigate radiation exposure, the weight of these aprons, coupled with the need for physicians to stand throughout the entire procedure, elevates the risk of orthopedic injuries. Finally, it is worth mentioning that fluoroscopic views are not constantly available, as the use of fluoroscopy is limited to crucial moments to minimize radiation exposure. Given this intermittent use, there is growing interest in exploring continuous imaging modalities that do not rely on radiation to provide visual feedback for catheter steering.

1.4 Emergence of robotic assistance

Robotically assisted endoluminal interventions involve combining robotic technology with non-motorized surgical and interventional tools. Figure 1.3 provides an overview of commercially available robotic systems for endoluminal procedures. A distinction is made between three specialties: gastroenterology, respiratory, and cardiovascular. In robot-assisted catheterization, the use of sophisticated actuation technologies allows robotic systems to be more precise, thus decreasing the likelihood of unintentional harm to the vasculature. Furthermore, by teleoperation, robotic

Table 1.1: Robotic Catheter Systems Approved by the FDA and Relevant Regulatory Authorities Worldwide

Products	Clinical uses	Main characteristics
Sensei X Robotic Catheter System (Hansen/Auris Medical Inc., CA, US)	electrophysiology procedures, such as treatment of cardiac arrhythmias	The Artisan control catheter has two robotically controlled segments which provides six degrees of freedom and 270 degrees of bend articulation.
Magellan catheter system (Auris Surgical Robotics, CA, US)	a diverse range of cardiovascular disease treatment procedures.	A robotic catheter system that can be operated remotely, based mainly on the technology of the original Sensei™ platform, but with significant improvements. The system consists of a 6-F leader catheter. This device offers improved maneuverability with 180-degree multi-directional articulation and a 9.5 F outer sheath that provides an extra 90-degree multi-directional articulation.
Corindus CorPath 200 (Siemens Healthineers, Germany)	precise positioning of coronary guidewires and balloon/stent for Percutaneous Coronary Intervention (PCI) procedures.	The CorPath system features an articulated arm, mounted to the procedure table, with a sophisticated robotic drive overlay. The guidewire or guide catheter is rotated in precise 30-degree increments and is inserted in one millimeter increments.
R-One robotic catheter (Robocath, France)	interventional cardiology	The R-One system comprises a command unit positioned in the cathlab and a robotic unit comprising the robot and its articulated support arm. The system is an open platform and is compatible with market leading devices (guidewires, stents / balloons, imaging systems, etc.)
Amigo® Remote Catheter System (Catheter Precision Inc, US)	electrophysiology (EP) and ablation procedure	The system is compatible with a wide range of standard EP catheters, which can be mounted to the robotic arm using a specialized attachment.
The Niobe ® Magnetic Navigation System (Stereotaxis Inc., US)	electrophysiology procedure	The system is composed of two robotically-controlled magnets positioned adjacent to the table. Physicians can manipulate the surrounding magnetic field via an intuitive computer interface, thereby allowing them precise control to direct and steer a catheter that has a magnet embedded at its tip.
LIBERTY® Robotic System (Microbot Medical Inc., US)	interventional cardiology	First ever single-use endovascular robotic system crafted to simplify endovascular procedures, uniquely designed to be compatible with off-the-shelf instruments, can be employed across various service locations.

systems mitigate the radiation exposure of clinicians. They also offer a more ergonomic operating environment, which reduces the fatigue of the physician.

Table 1.1 provides an overview of the main robotic catheter systems that have received clearance from the FDA or equivalent authorities. All these systems adopt teleoperation, where clinicians manipulate control input devices, such as a joystick, in a separate control room. In doing so, the robot, positioned next to the patient, performs the necessary actions.

In most systems, the catheter driver is attached to a robotic arm, allowing for adjustments of the insertion angle to the patient's body. The robotic device itself manages the insertion/retraction, rotation, and bending of the catheter. Some systems, including the Amigo® and R-One®, are compatible with various commercially available catheters. On the contrary, others, such as the Sensei X® Robotic Catheter System[1], are designed for specific types of catheters. In terms of application, systems such as Sensei X®, Amigo®, and Niobe® are used predominantly for electrophysiology and ablation procedures. Systems such as Magellan®, Corindus®, CorPath®, and R-One® demonstrate versatility, accommodating a wide range of devices and procedures. Lastly, when it comes to actuation technologies, most robotic catheter systems, with the exception of the Niobe®, rely on a cable-driven approach to bend the tip of the catheter. The Niobe catheter system, on the other hand, operates through magnetic actuation. Most of the catheter systems mentioned above share a similar design philosophy. However, the LIBERTY® system from Microbot Medical Inc. stands out as the world's first disposable catheter system. This reduces preparation efforts and the expense involved in establishing a sterilized environment. In terms of design, the external structure of the LIBERTY® robotic system is compact, with the entire device being approximately the size of an adult's palm. This simplistic design approach largely reduces manufacturing costs. In terms of configuration, the LIBERTY® system comprises a control system and an operating handle. The control system can be attached to the patient's thigh, thus a cumbersome robotic arm to support the catheter driver is not needed, while the physician uses a gaming joystick to manipulate guidewires and catheters during the procedure. This makes the catheterization straightforward.

Nevertheless, these systems are not without limitations. One primary limitation is that these systems are mainly used for simpler interventions, but rarely for more complex ones (such as coronary chronic total occlusions), where the highest mental load is expected from the clinician. Another limitation is the lack of autonomy. Currently, robotic catheter systems operate entirely under the control of a human operator, which means that the movements of the robotic catheter merely mimic those executed by the physician. As per the classification system for levels of autonomy in surgical robotics outlined in [25], this scenario qualifies as Level 0 autonomy. However, advancements to level 1 with surgical assistance or level 2 task automation are anticipated and remain areas of ongoing research. Also, as introduced in Section 1.3, the incorporation of robotic assistance and teleoperation

---

[1] The Sensei X robotic system is no longer available on the market.

presents additional challenges, such as the lack of real-time haptic feedback and poorer hand-eye coordination, compared to the current practice where the physician can directly manipulate the catheter. In addition, in terms of visual feedback, these systems still largely rely on traditional fluoroscopy. Although physicians are shielded from radiation, patients and other clinicians involved in swapping various interventional instruments remain exposed to radiation. Using fluoroscopy also limits the visual feedback of physicians to 2D grayscale images. Furthermore, clinicians unfamiliar with robotic systems will need to overcome a characteristic learning curve [26].

1.5 Application of ML techniques

ML, a subfield of artificial intelligence, has a long history dating back to the middle of the 20th century. The term “machine learning” was built on the idea that machines can learn from data. During the 1960s and 1970s, the area of study broadened, with scientists investigating a variety of learning algorithms and theoretical foundations. The rise in computing power substantially boosted the growth of ML. This led to the development of more sophisticated models and algorithms. The exponential growth of the Internet in the 21st century significantly enhanced the field of ML by providing an abundance of data, which is essential for training algorithms.

ML has been adopted in various fields today. A notable area is surgery, where ML has been used for tasks such as image analysis [27], diagnosis support [28], and treatment planning [29]. Recently, this trend has started to extend into the field of Flexible Surgical and Interventional Robots (FSIR), encompassing instruments such as catheters and guidewires. Incorporating ML techniques alongside robotic assistance can open up many opportunities to address various challenges in current clinical practice.

1.6 Motivation of this thesis

Despite advancements in introducing robotic assistance into the current catheterization procedure, a significant gap remains. The commercially available robotic catheter systems mentioned above are designed or only approved by relevant authorities for relatively simple interventions (such as peripheral vascular interventions). Their ability to reliably perform complex procedures, such as addressing coronary chronic total occlusions, has not been demonstrated. This explain why existing catheterization procedures remain predominantly manual. Current commercial robotic catheter systems are not intelligent and foresee no automation, instead they are solely “teleoperated”. The catheter driver that steers the catheters merely replicates the motions input by the operator. Consequently, the control precision and safety of these procedures are a.o. highly dependent on the level of expertise of the operating physician. Furthermore, the change from direct catheter manipulation to teleoperation, the need to manage various joysticks and buttons for catheter control simultaneously, coupled with the absence of force feedback in teleoperation, make a considerable learning curve for physicians [26, 30].

Moreover, hysteresis continues to be a challenge even with the introduction of robotic assistance, and may even become worse due to the additional non-linearities introduced by the employed actuation technologies in the robotic solutions. Regarding visual feedback, most commercially available endovascular robotic systems continue to depend on fluoroscopy. This reliance raises significant concerns about radiation exposure for both patients and clinicians, particularly those who handle instrument swaps during interventions. Furthermore, the fluoroscopy image, being a 2D grayscale representation, significantly limits clinicians' ability to perceive depth information.

1.7 Overarching goal of the thesis and sub-objectives

Guided by the above motivations, the overarching goal of this thesis is

> *to enhance the precision and intuitive control of robotic catheters for safe navigation in complex environments.*

The overarching goal of this thesis is comprised of several distinct but interconnected sub-objectives.

- **Sub-objective 1:** identification of the latest trends of ML techniques in FSIR.

  Understanding the latest applications of ML techniques in the FSIR field is crucial for leveraging ML methods to tackle key challenges of catheter control in complex environments. ML offers an alternative approach to model and control robotic catheters compard to traditional analytical models. Through accurate catheter modeling, precise catheter control could be achieved, which is essential for navigating through complex vascular structures. Additionally, ML could also lead to an elevation of the level of autonomy in catheter steering. By learning from a vast array of procedural data, ML could automate certain aspects of catheter control, reducing the reliance on manual manipulation and potentially leading to more efficient and safer procedures.

- **Sub-objective 2:** investigation of the most effective interactive modality for tele-operated endovascular intervention.

  Teleoperation is anticipated to be crucial in the long term, particularly in complex environments such as navigation in the coronary arteries. Exploring intuitive interaction modes for teleoperation is therefore of great significance. Having an intuitive mode for catheter steering, coupled with reliable visual feedback, could also contribute to safer interventions.

- **Sub-objective 3:** determination of the 3D catheter shape using a non-radiative method.

  Reconstructing the 3D shape of catheters using non-radiative methods reduces radiation exposure for both clinicians and patients, thereby making interventions safer. In addition, especially in complex environments such as vasculatures, having real-time knowledge of the catheter's 3D shape and motion is crucial for effective catheter control.

- **Sub-objective 4:** identification and compensation of non-linearities in catheters.

  Understanding and compensating for the nonlinear phenomena in catheter systems can significantly enhance the controllability of catheters. First, in teleoperation cases, it improves the intuitiveness of catheter control. Without such compensation, users would have to adapt and learn to offset these nonlinear behaviors themselves. Second, in the context of automatic control, compensating for non-linearities allows for the derivation of a precise kinematic model of the catheter, thereby enhancing the control accuracy of the catheter. Both these aspects could be crucial to achieve precise and safe interventions in tortuous vessels.

- **Sub-objective 5:** elevation of the level of autonomy in robotic catheters, along with its demonstration in an *in-vitro* environment.

  Autonomous catheterization implies that the catheter can navigate autonomously within the blood vessels and perform the corresponding tasks, with physicians merely monitoring the entire process. Achieving fully autonomous catheter navigation in *in-vivo* environment poses significant challenges. Recognizing these difficulties, this thesis sets a more realistic and achievable goal: demonstrating autonomous or semi-autonomous catheter navigation in an *in-vitro* experiment. This approach not only stems from an understanding of current technological challenges, but also aims to incrementally advance the level of autonomy of catheter control, laying a solid foundation for future applications in more complex environments.

  (Semi-)autonomous control could potentially lead to catheter motion with consistent precision, thus mitigating performance variations among physicians with different levels of experience. Second, autonomy in catheter control could reduce the physical burden on clinicians, allowing them to focus more on high-level tasks such as decision-making and minimizing performance degradation due to fatigue. In addition, conducting experiments in an *in-vitro* environment can validate the robustness and safety of the control algorithms, laying the foundation for further validations. All of these factors could potentially contribute to precise catheter control and ensure safer interventions.

Above five sub-objectives could together address these key limitations, as detailed below: 1) limited intuitiveness and assistance in catheter control based on teleoperation, further challenged by the 2D grayscale nature of fluoroscopy imaging; 2) the reliance on radiative fluoroscopy for the reconstruction of the catheter's 3D shape; 3) the presence of nonlinearities in catheter systems, which reduce precision and intuitiveness in control; and 4) no level of autonomy in both current practices and robotic solutions.

1.8 Thesis outline

Figure 1.4 offers a diagrammatic representation of this thesis, illustrating the primary sub-objectives alongside their respective chapters. The contents of each chapter are as follows:

Chapter 1 reviews the history of surgery, introduces knowledge/background related to cardiovascular diseases, and their current treatment methods. This chapter also analyzes the limitations and challenges of existing catheterization methods. In response to these challenges, robotic assistance and DL techniques are proposed. Finally, this chapter specifically outlines the overarching goal of this thesis, the sub-objectives and the detailed content of each subsequent chapter.

Chapter 2 provides a comprehensive survey of ML techniques, including traditional ML methods and modern DL approaches, in flexible surgical and interventional robots. The chapter summarizes the different roles of ML/DL techniques in the perception, modeling, navigation, and control of the FSIR. The trends, popularity, and strengths of applying ML/DL algorithms in such robotic systems are demonstrated and analyzed. A discussion on the potential pitfalls of ML/DL in the context of FSIR is included. Chapter 2 provides insight into selecting the most appropriate algorithm for specific tasks, which inspires the algorithm selection for the work in the subsequent chapters.

Chapter 3 conducts a user study exploring various interactive modes derived from multiple control input devices and visualization devices. An AR interface for a Head-Mounted Display (HMD) is designed for enhanced visual feedback and teleoperation. Joystick-based teleoperation to steer a robotic catheter system and demonstration of visual feedback using a 2D monitor were implemented for comparison. A user study is described that involves nine participants with diverse experience levels in AR-HMDs, gaming, and steerable catheters. A detailed performance analysis using both subjective and objective evaluation metrics is included.

Chapter 4 introduces a new DL method for improved shape sensing of continuum robots based on the multi-core FBG sensor. The suggested method enables shape sensing using a single Multi-Core Fiber (MCF) positioned off-center. This facilitates miniaturization of the catheter while freeing up the central channel for additional uses. A 2-DOF benchtop fluidic-driven catheter system is built to validate the

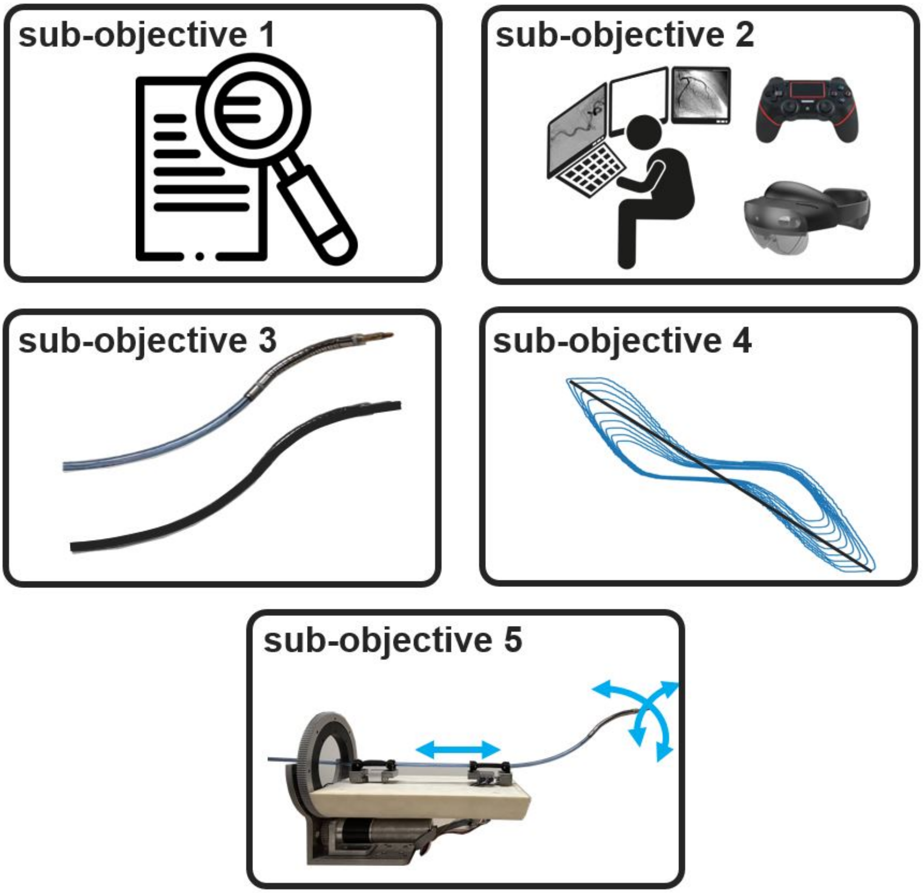


Figure 1.4: Illustration of the research sub-objectives of this thesis. **Sub-objective 1:** identification of the latest trends of ML techniques in FSIR (Chapter 2). **Sub-objective 2:** investigation of the most effective interactive modality for teleoperated endovascular intervention (Chapter 3). **Sub-objective 3:** determination of the 3D catheter shape using a non-radiative method (Chapter 4). **Sub-objective 4:** identification and compensation of non-linearities in catheters (Chapter 5 and Chapter 6). **Sub-objective 5:** improvement of the level of autonomy in robotic catheters, along with its demonstration in an *in-vitro* environment (Chapter 6).

proposed ANN. Ablation studies have been conducted to enhance the comprehension of the ANN, thereby augmenting its interpretability. The proposed method offers precise shape reconstruction of a robotic catheter without resorting to radiation-based imaging techniques, such as fluoroscopy, while accurate representation of catheter shape offers new valuable information for catheter control.

Chapter 5 addresses the nonlinear phenomena in catheter systems. A Recurrent Neural Network (RNN) is used to model the hysteresis of a robotic catheter. This approach aligns with the inherent characteristics of hysteresis and RNN, where the current system output is influenced not just by the present input but also by historical inputs. For a comprehensive assessment, both an analytical model and a traditional ML approach are compared. The proposed method was further validated with a reconstruction task. A forward-looking fiber with optical coherence tomography was used to scan an artificial environment. The method can predict the response of the catheter tip solely based on the proximal input pressure while being robust to hysteresis. This reduces the challenge of embedding sensors at the tip of the catheter, which is challenging due to spatial constraints and sterilization needs.

Chapter 6 introduces a DL approach to compensate for the hysteresis in catheters. The trained ANN also works as an effective position controller, ensuring precise steering of the catheter in the free space. Additionally, a compliant motion control method is proposed to automatically reduce the contact force between the catheter and the environment. Remarkably, the method does not require the installation of a force sensor at the catheter tip.

Chapter 7 explores a plan to valorize the findings of this thesis. It begins by summarizing key research results, followed by an in-depth analysis of the prevailing medical market. Building on this groundwork, the chapter details the crucial strategies for valorizing the DL algorithms and datasets derived from this doctoral research. In addition, the chapter proposes preliminary methods for licensing these algorithms and datasets.

Chapter 8 presents an evaluation of the research undertaken. Initially, it offers a concise summary of the thesis's primary objectives and key discoveries. Subsequently, the chapter delves into a critical examination of the research's limitations, providing insights into areas where the study could be further refined. The chapter concludes with recommendations for future research, identifying specific domains that could benefit from and extend the findings of this thesis. Additionally, it underscores the persistent imperative for ongoing exploration and advancement within this field.

1.9 Collaborations

In this thesis, the primary emphasis is on leveraging DL and augmented reality technologies to address various challenges in the sensing and control of robotic catheters. The authors express deep gratitude for the collaborations with fellow researchers, which have led to joint publications. The specific contributions of the author and other researchers are clarified as follows:

In Chapter 2, both the author and Renchi Zhang made equivalent contributions to the manuscript's composition. While the author focused primarily on surgical robotics aspects, Renchi Zhang provided insights from the ML perspective. In Chapter 3, the author contributed to the setup of the experimental system, designed the user study, and conducted an analysis of the experimental results. Zhen Li contributed jointly to the setup of the experimental system and the processing of the experimental data. In Chapter 4, the author was responsible for the design, manufacturing and assembly of the experimental setup. Furthermore, the author implemented the neural network approach along with the training and validation of the neural network's performance. Xuan Thao Ha is primarily responsible for data collection and validating the proposed auto-calibration method. In Chapter 5, the author proposed the idea of using LSTM to model hysteresis, performed the neural network training, and prepared and conducted the environmental reconstruction experiment based on LSTM and OCT. Yao Zhang contributed to the preliminary validation of the LSTM and jointly handled the data processing. In Chapter 6, the author conceived the idea of employing LSTM for catheter control [31] and implemented the training and validation of the LSTM. Moreover, the author collaborated in implementing the iDRDPI model, set up and jointly conducted the phantom study. Xuan Thao Ha validated the compliant motion controller, jointly conducted in the phantom study, and analyzed the results.

Chapter 2

Machine learning in flexible and interventional robots

This chapter is based on the submitted journal article:

D. Wu*, R. Zhang*, A. Pore, D. Dall'Alba, X. T. Ha, Z. Li, Y. Zhang, F. Gonzalez, M. Ourak,W. Kowalczyk, E. De Momi, A. Casals, J. Dankelman, J. Kober, A. Menciassi, P. Fiorini, E. Vander Poorten, "A Review on Machine Learning in Flexible Surgical and Interventional Robots: Where We Are and Where We Are Going".

Minimally Invasive Procedures (MIPs) have emerged as an alternative to more invasive surgical approaches, offering patient benefits such as smaller incisions, less pain, and shorter hospital stay. In one class of MIPs, where natural body lumens are used to access deeper anatomical locations, Flexible Surgical and Interventional Robots (FSIRs) such as catheters and endoscopes are widely used. Due to their flexible and compliant nature, FSIRs can be inserted via natural orifices or small incisions, then moved towards hard-to-reach targets to perform interventional tasks. However, existing FSIRs suffer from control and navigation limitations due to their non-linear behavior and the complexity of path planning in intricate lumens where modeling interaction disturbances is very complex. Rapid advances in Machine Learning (ML) have supported the wide adoption of ML methods in FSIRs. This chapter provides an overview of these efforts by first introducing a classification of existing ML algorithms, including traditional ML methods and modern deep learning approaches, commonly used in FSIRs in Section 2.2. Next, the use of ML algorithms is surveyed per sub-domain, namely for perception, modeling, control, and navigation, in Section 2.4. Trends, popularity, strengths, and/or limitations of different ML algorithms are analyzed. The different roles that ML plays among tasks are investigated and described. Finally, discussions are conducted on the limitations and prospects of ML in MIPs in Section 2.5.

2.1 Introduction

The first chapter introduced that ML techniques have been used in the field of surgery, and our aspiration to apply ML to robotic catheter control. To address sub-objective 1 introduced in Section 1.7, a comprehensive review of Machine Learning (ML) techniques used in Flexible Surgical and Interventional Robots (FSIRs) is conducted. This review could help determine the most suitable ML approaches to address the challenges of robotic catheter control, particularly in complex environments. ML techniques can be employed to enhance precision and intuitiveness in catheter control and to reduce its interaction force with the environment. This leads to safer interventions, aligning with the overarching goal of the thesis. This chapter broadens the scope from DL to all ML techniques, from cardiac catheterization to all Minimally Invasive Procedures (MIPs), and from catheters to include all types of FSIRs. Given the similarities shared among various types of FSIRs, the broadened scope of the review is designed to provide more information and inspiration that may be transferred and useful between disciplines.

MIPs have revolutionized the field of surgery since their advent in the 1980s [32]. Small wounds, shorter recovery time, and improved cosmesis [33], all of these benefits favored the adoption of MIPs with respect to more invasive open surgical procedures over the past decades. Currently, MIPs are adopted in different medical specialties and have formed different subtypes, such as natural orifice procedure (e.g., colonoscopy, ureteroscopy), endovascular catheterization, and minimally invasive brain surgery [34]. In MIPs, one/multiple small incisions are made, or natural openings are used to access the body and reach the target areas of interest. Commonly, MIP access ports are narrow and the lumen can be tortuous, so it is advantageous to use snake-like instruments, such as catheters and flexible endoscopes, since they can reach deeper areas in the lumen. Currently, the (steerable) instruments commonly used in clinical practice are controlled manually. This makes MIPs highly skill-intensive.

FSIRs have emerged to overcome the limitations of current (steerable) clinical instruments. They try to make these approaches less complex for clinicians. FSIRs refer to those snake-like steerable instruments of which one or more degrees of freedom (DOF) can be commanded in a computer-controller fashion. FSIRs take advantage of progress in different robotic fields: actuation, sensing, manipulation, and control. Different design and manufacturing approaches have been used to implement FSIRs, thus making the structure or embodiment of FSIRs different. In this chapter, FSIRs refer to all snake-like robots used for MIPs: single/multi-backbone “continuum” robots [35], concentric tube robots (CTRs) [36], multi-joint flexible robots [37], and flexible needles [38]. FSIRs has been shown to be advantageous in accessing areas that are difficult to reach within the intricate and delicate internal anatomy of the human body [39, 40, 41]. Moreover, they could offer a higher level of autonomy, intuitiveness, and precision than passive flexible instruments [42]. The use of FSIRs could potentially reduce mental burden during both challenging and routine tasks, so that clinicians can focus more on important aspects or other critical tasks.

2.1.1 Workflow and tasks of FSIR systems

The workflow of robot-assisted surgery can be broken down into three distinct stages: sensing, planning, and execution [43]. These phases correspond to three tasks, i.e., perception, navigation, and control. These tasks, together with the modeling, constitute the four primary FSIRs tasks discussed in this work and are visualized in Fig. 2.1.

The perception of FSIRs is based on intraoperative sensors that can be integrated directly into FSIRs or external to the patient. Based on the data collected by the sensors, FSIRs can acquire information about their current state, the anatomical environment, and the progression of the current task or the general procedure. During the navigation phase, the various information perceived, combined with both the task goal and the current status of the procedure, is used by the planner to plan the robot's motions. As for the control, the controller computes the commands according to the previously planned motions from the navigation, and sends them to the actuators. Accurate modeling of the robot system makes the input/output relations clear, thus is crucial for control accuracy.

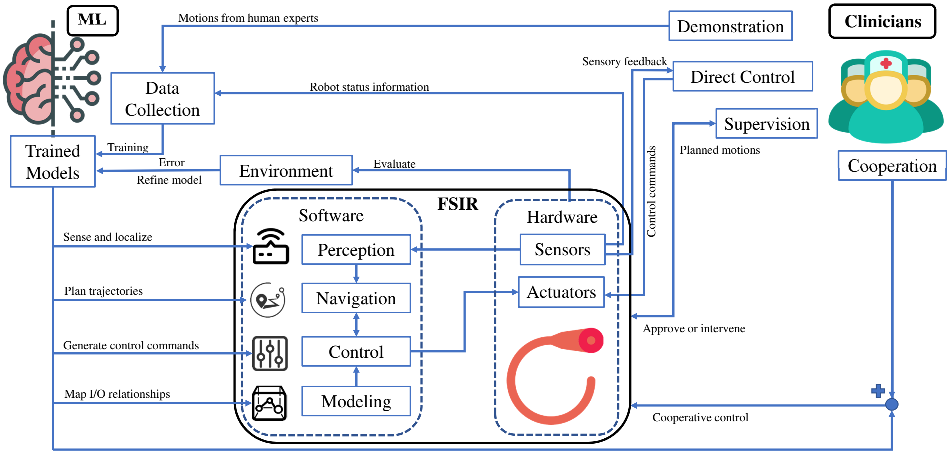


Figure 2.1: An overview of the interactions between ML algorithms, clinicians, and FSIRs in robotically assisted MIPs. Clinicians can control FSIRs directly with sensory feedback or cooperatively with the assistance of ML. The ML can be trained by the data collected from the FSIRs, or by the demonstration data of expert clinicians. To improve safety, clinicians need to supervise the whole workflow when ML is used.

2.1.2 Motivation for machine learning in FSIR

An increasing number of researchers have started investigating methods based on ML to overcome the existing challenges in FSIRs. The objective of ML is to enable a system to gain knowledge by "learning" from examples or experience (training data), and then to make predictions or represent regularities from the acquired knowledge. The learned knowledge could be used in different tasks, such as recognizing the hidden patterns in data, or generating a mapping between input and output, depending on the applications.

Table 2.1: General pros and cons of ML approaches

Pros	Cons
• No need to understand the underlying physics. No simplifications and assumptions of the system required	• Involving more uncertainty due to the lack of interpretability (cause of a decision)
• Flexible transferability to different tasks with few expert knowledge	• Significant computational resources required
• Increasing the level of autonomy and cognitive capabilities of a system	• Model performance depending crucially on the quality and amount of data, or the quality of manual feature engineering

In the context of FSIRs, the use of ML methods could avoid complicated modeling and its associated cumbersome parameter identification procedures. ML can capture the underlying patterns from the data generated by the system [44]. Additionally, ML-based approaches are observed to have good generalization capability, which allows them to be possibly transferred among different types of robot and different anatomical environments [45, 46, 47]. ML could potentially increase the level of autonomy and cognitive abilities of FSIRs [48]. Consequently, applying ML to FSIRs could potentially free the surgeon or interventionist from routine tasks and allow increased focus on higher-level tasks such as decision-making. ML-assisted FSIRs could also potentially lead to faster execution of routine tasks or lead to higher precision and smoother flow or trajectories compared to manual operation by human experts [48]. In the context of the FSIRs, ML approaches are generally considered to have pros and cons [48, 49, 50] as shown in Table 2.1. The roles of ML algorithms, clinicians and FSIRs are summarized in Fig. 2.1. The figure shows that data to train ML algorithms could be collected from clinicians (e.g. human motions) or FSIRs (e.g., actuators' status, shape and pose of the FSIRs). In recent years, the data collected from FSIRs systems has experienced a progressive growth. This has supported the adoption of data-driven approaches such as ML. The trained ML algorithms are used in perception, navigation, control and modeling of FSIRs. Clinicians may provide demonstrations for training data [51, 52], give direct control to FSIRs, supervise FSIRs [53, 54], or collaborate with FSIRs in different phases of MIPs [55]. FSIRs can reduce the burden of clinicians on both mental / cognitive and physical aspects [48].

2.1.3 Summarized contributions of this review

Over the past years, some review articles have introduced specific aspects of FSIRs related to fabrication[42], actuation [34, 39, 42, 56], sensing [56], modeling[39, 41, 42, 56], imaging and navigation[40, 57], and control[39, 41, 42, 56]. Specific ML techniques are partially discussed in some of these reviews, but do not constitute their main contents. Some other reviews focus on the usage of ML techniques across the entire surgical robotics field [43, 48], while their discussion of FSIRs is sparse. ML used in soft robots has been reviewed in [58], but this work considers generic contexts without a specific link to clinical applications. Moreover, some of these reviews were published five or more years ago [39, 40, 43, 48, 57]. Given that the ML techniques have gained popularity in recent years, it makes sense to revisit the conducted works to capture the most recent trends.

To the best of the author's knowledge, no previous review has been performed for the ML techniques specifically used on FSIRs. This chapter tries to fill this gap by providing an overview of ML, including traditional approaches and modern deep learning approaches, and discussing general approaches to leverage ML in various FSIRs tasks. It discusses the limitations and perspectives of ML algorithms in FSIRs. The chapter is structured as follows: Section 2.2 introduces ML and their classifications; Section 2.3 describes the bibliography search method and the selection criteria; the ML methods used in different FSIRs tasks, i.e. perception, modeling, control, and navigation, are summarized and compared in Section 2.4. Section 2.5 discusses the current challenges in ML, while Section 2.6 concludes the whole chapter.

2.2 Machine learning techniques

ML is described as "*programming computers to optimize a performance criterion using example data or past experience*" [59]. Current ML algorithms can be divided into three main categories: supervised learning, unsupervised learning, and reinforcement learning (RL) [60]. Supervised learning learns the mapping from input data to ground truth (labels) using annotated datasets, while unsupervised learning learns to discover regularities solely from the input data. Rather than learning from input data, RL trains an agent while interacting with the environment using a system of reward and punishment, thus learning from the agent's experience[61]. The presented conventional classification is clear and concise, but is not instructive enough to differentiate and analyze the characteristics of traditional ML methods and the increasingly popular DL method. Thus, we propose a two-dimensional (2D) classification method (see Fig. 2.2) for ML, as this representation can provide an additional perspective on algorithm choice for FSIRs applications. The conventional classification above (supervised learning, unsupervised learning, and RL) is preserved as the first dimension in our proposed 2D classification.

However, supervised learning and unsupervised learning are merged into a single category. Furthermore, in our 2D classification, ML algorithms are categorized into two distinct groups: traditional ML and DL approaches.

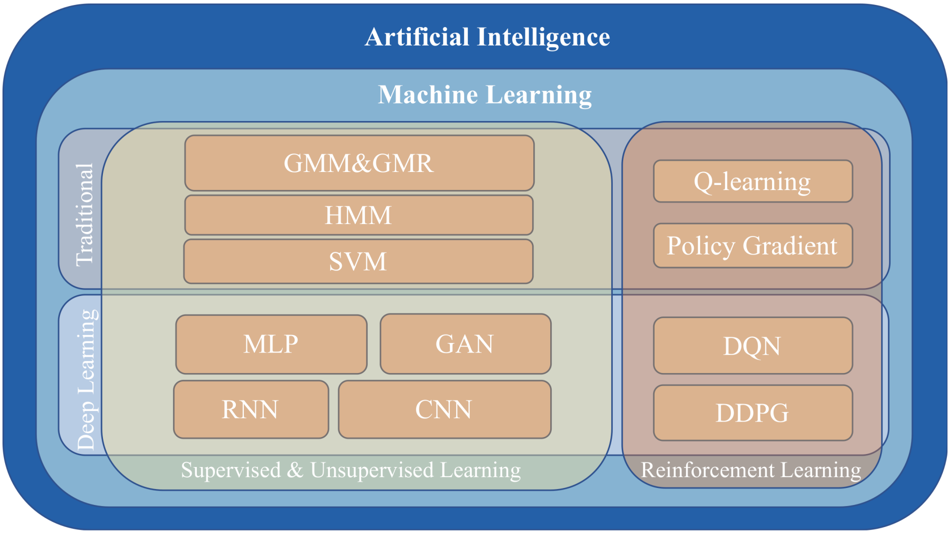


Figure 2.2: A Venn diagram of AI techniques applied on FSIRs: ML is the dominant subdiscipline of AI. ML can be divided into traditional ML and DL. Several specific ML algorithms often used in the FSIRs are also indicated in the figure. The readers could refer to Table 2.2 for the abbreviations of the different ML algorithms.

In this review, traditional ML refers to ML algorithms that do not rely on deep Artificial Neural Network (ANN). Traditional ML includes many algorithms that have been studied for a long time. In this decade, the use of deep ANN has improved the state-of-the-art in many domains such as visual object detection, natural language processing, speech recognition, among others. [62]. DL is also embraced by the FSIRs community these days. However, at the same time, traditional ML approaches maintain their significance in a range of FSIRs tasks. Thus, demonstrating the applications of traditional ML and DL separately in different tasks is helpful for algorithm selection. The remaining of this section details the two dimensions of our classification method.

2.2.1 Supervised, unsupervised and reinforcement learning

In the classification illustrated in Fig. 2.2, supervised and unsupervised methods are merged even though they are distinct categories of ML techniques. The decision to unify these methods arises from the observation that various ML techniques applied in FSIRs are capable of solving both supervised and unsupervised problems

Table 2.2: Abbreviations List of Terminology (in alphabetical order within each ML category)

Category	Terminology	Abbr.
Deep Learning (DL)	Artificial Neural Network	ANN
	Convolutional Neural Network	CNN
	Generative Adversarial Network	GAN
	Long Short-Term Memory	LSTM
	Multilayer Perceptron	MLP
	Recurrent Neural Network	RNN
Reinforcement Learning (RL)	Deep Deterministic Policy Gradient	DDPG
	Deep Q-Networks	DQN
	Deep Reinforcement Learning	DRL
Traditional Machine Learning (ML)	Gaussian Mixture Model	GMM
	Gaussian Mixture Regression	GMR
	Gaussian Process Regression	GPR
	Hidden Markov Model	HMM
	k-Nearest Neighbors	kNN
	k-Nearest Neighbors Regression	kNNR
	Markov Decision Process	MDP
	Random Forest	RF
	Support Vector Machine	SVM
	Extreme Learning Machine	ELM

when provided with a well-curated dataset. Supervised learning requires input data with its corresponding annotations, which can be discrete (for classification) or continuous (for regression). Supervised learning aims to identify the relationship between inputs and labels, thereby enabling accurate predictions when faced with unfamiliar input data. In FSIRs, supervised learning applications range from understanding the kinematics of FSIRs[63], to predicting control commands [64], or determining the status of FSIRs [65], among others. Unsupervised learning, which identifies patterns or correlations within the data, does not require labeled data. Consequently, unsupervised learning methods can be employed to discover feasible FSIRs trajectories in Learning from Demonstration (LfD) tasks [66], or to estimate depth and motion configurations from endoscopic image data [67]. The effectiveness of supervised and unsupervised approaches is closely tied to the quality of the training dataset. However, curating a high-quality dataset can be a complex, time-consuming, and human-dependent task, particularly when generating annotations for supervised learning.

The third category of ML approaches, RL, is ideal for situations in which the autonomous system can continuously learn from its interaction with the environment by trial-and-error. RL learns the policy, i.e., a series of correct actions that achieve the goal in defined situations [61]. At a given moment, the agent takes an action that leads to a new environmental state. The agent is given either a reward or a punishment (namely a negative reward) from the environment after taking an action. The training aims to maximize the numerical cumulative rewards received by the agent. For a considerable period of time, RL has been employed in tasks related to planning [68] and control [69] (if the control task is relatively easy to execute), as the robot can be continuously trained to adapt to changes in the environment. RL trains agents without the need for labeled datasets, which are typically essential for supervised learning techniques.

2.2.2 Traditional ML versus DL

As introduced in the previous section, traditional ML is a group of ML algorithms that are not based on deep ANN. Contrary to this, DL relies on deep ANN. The performance optimization of traditional ML algorithms often requires the effort of feature engineering (i.e. transforming raw data into informative and representative values and selecting these values for learning tasks) or domain-specific knowledge to preprocess training data [70] (see Fig. 2.3). When appropriately tuned, traditional ML algorithms can often deliver relatively solid performance, especially when working with limited or lower-dimensional data sets that require minimal preprocessing and feature selection, or when computational resources are limited [71]. Moreover, some traditional ML algorithms, including linear models and decision trees, are recognized for their high interpretability. Their clear mathematical and statistical foundations make them easily comprehensible for humans. These models can also serve to clarify the working mechanism of other traditional ML and deep ANN models [72].

The most common traditional ML algorithm employed in the FSIRs tasks is the Gaussian mixture models (GMM) and the Gaussian mixture regression (GMR) approach, according to the statistics shown in Fig. 2.4. GMM assumes that the data distribution of the input follows a combination of multiple Gaussian distributions, each with its own parameters [60]. Therefore, GMM could represent the joint density of the data. For each Gaussian model, its conditional density and regression function can be derived. GMR utilizes these GMM regression functions to make predictions for new inputs [73]. The GMM-GMR approach is often used in LfD tasks in FSIRs [64, 66, 74, 75, 76]. It can extract statistical characteristics from noisy demonstration data (that is, typical when acquiring data from real FSIRs) and predict continuous control commands over time [66]. Based on the statistics in Fig. 2.4, another popular traditional ML algorithm, the Support Vector Machine (SVM), is also used in FSIRs to tackle classification tasks [77, 78].

DL is implemented based on deep ANN. ANN originated from the idea of mimicking the working principle of the human brain, yet it is still far from the ability of human brains at the current stage. A neuron is a basic computational unit of an ANN. It is connected to a number of inputs and provides an output. To generate an output, a neuron multiplies the inputs by the corresponding weights and then

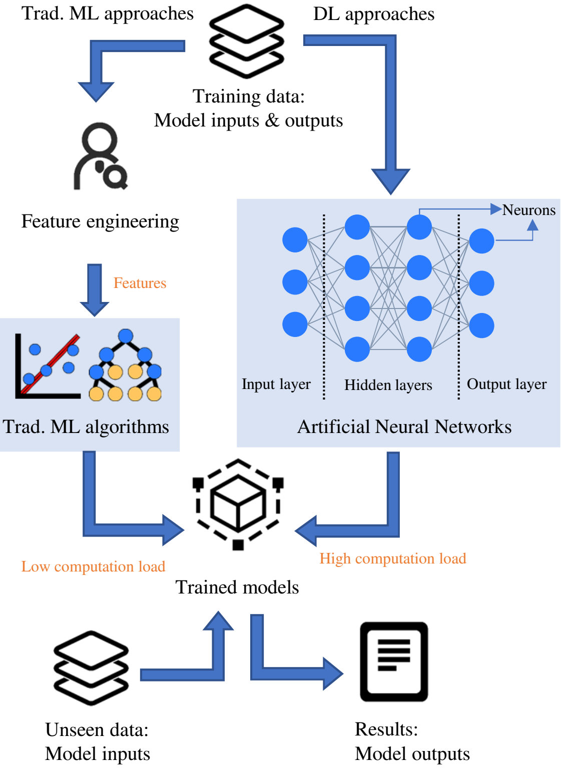


Figure 2.3: The comparison between the workflow of traditional ML and DL in supervised learning tasks.

applies a non-linear activation function. Multiple neurons at the same level of connection form a layer. When the layers are connected to other layers, an ANN is generated (see ANN in Fig. 2.3). The input layer of an ANN is responsible for receiving the input data, and the output layer provides the prediction of ANN. The layers located between the input layer and the output layer are hidden layers.

In the 1980s-1990s, some classical algorithms, which have had a lasting impact on the field, such as Long Short-Term Memory (LSTM) [79] and Convolution Neural Network (CNN) [80] were proposed. At the same time, backpropagation was proposed and used to train deep ANN [81]. In recent years, rapid advances in storage, computing power, and advanced software libraries have increased DL. Advancements in hardware have also facilitated a considerable reduction in training time of ANN [82]. At the same time, the explosion of big data offers the possibility of training ANN from huge datasets. Moreover, emerging optimization methods (e.g. Adam), activation functions (e.g. ReLU), and regularization methods (e.g. Dropout) are also driving the development of DL. Furthermore, the emergence

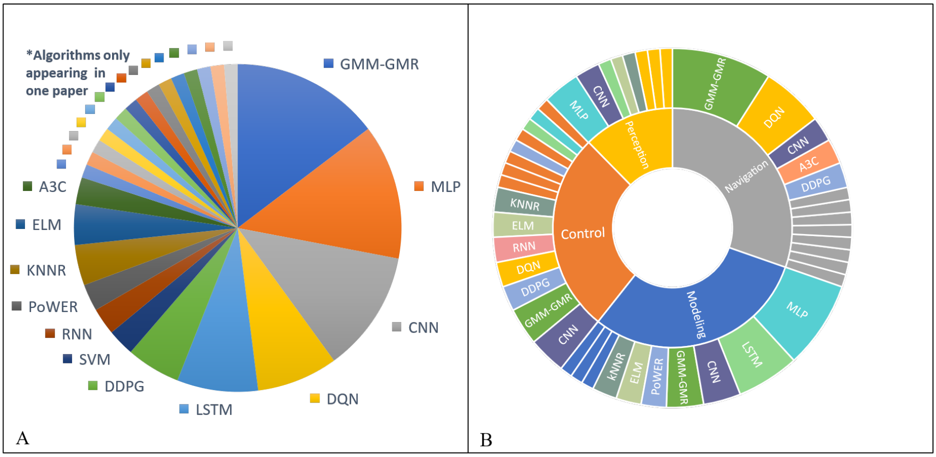


Figure 2.4: The prevalence of various ML algorithms in search results. (A) The relative amounts of each algorithm in all outcomes. (B) The relative amounts of each algorithm used in various tasks are shown. *Algorithms include UDQL, SARSA, Q-learning, MDP, MA fuzzy Q-learning, LWPR, K-means, HMM, GAIL, BNN, AutoEncoder, Gaussian Process Regression, Gradient Boosting, and ANFIS.

of open source DL libraries (e.g., Tensorflow, Pytorch, Caffe, and Keras) further support the use and dissemination of DL-based approaches. All of these factors make DL accessible to researchers from various backgrounds [83]. This is also the reason why DL has become increasingly popular in engineering fields other than computer science.

DL avoids the need to craft particular features that extract the essence of the underlying data (see Fig. 2.3). Feature design is typically critical in traditional ML models, especially when dealing with high-dimensional data. Instead, DL tends to recognize such features automatically during training by adjusting a large number of ANN parameters. However, there exists the potential to improve the interpretability of DL models, thus addressing the increasing safety concerns associated with their application in high-risk fields such as healthcare [84]. There are different architectures of ANN among the DL algorithms, but only a limited set has been applied in FSIRs.

As indicated by the statistics shown in Fig. 2.4(A), the Multilayer Perceptron (MLP), which is also called the Feedforward Neural Network (FNN), is the most commonly employed ANN within the FSIRs domain. The widespread use of MLP can be attributed to its robust ability to model non-linear relationships prevalent in FSIRs, such as kinematic and dynamic modeling [63, 85, 86]. Furthermore, MLP's simplistic architecture lends itself to straightforward implementation and training.

CNNs are also often used with FSIRs. CNN is a type of ANN that contains convolutional layers. Convolutional layers contain filters/kernels used to extract features efficiently from the high-dimensional input data (e.g., images) by a convolution operation. CNN and its variants have demonstrated their strong capability to deal with high-dimensional data. Therefore, CNNs are frequently employed to process visual information from FSIRs in order to provide imaging guidance or forecast the current state (e.g., tip orientation, distal-end force) using data from multiple sensors [87, 88, 89].

RNN is another category of ANN that is often applied in FSIRs. Unlike MLP and CNN, which only consider current input during training, RNN can consider historical input to update their "memory" (i.e., hidden states) to influence its current outputs [90]. This structure allows the RNN to capture sequential information, such as time series data from the input. Within the field of FSIRs, RNNs can be used to tackle history-dependent phenomena such as hysteresis, or predict the distal end status of FSIRs based on proximal information [89, 91, 92].

Beyond the diverse ANN categories, Deep Reinforcement Learning (DRL) has emerged as an area that captures the interest of FSIRs researchers. Located at the intersection of DL and RL, DRL is one of the most rapidly evolving subjects in ML over the past decade, and frequently used to tackle real-world decision-making tasks. Its efficacy in reaching human-like intelligence has been demonstrated in a variety of games including Go, Chess, and Dota2 [93, 94, 95]. Within diverse DRL algorithms, ANN assumes varying roles. For example, in Deep Q-Networks (DQN) [96], ANN functions as a mapping tool, identifying the relationship between the input actions coupled with states, and the corresponding reward values. Recent research in FSIRs has exploited DRL algorithms such as DQN and Deep Deterministic Policy Gradient (DDPG) [97] to optimize trajectories, plan motion, and control FSIRs [54, 69, 98, 99, 100].

2.3 Literature search methodology

For this review, the Scopus database was used to perform a thorough literature search. Scopus is favored due to its extensive coverage of peer-reviewed literature, which covers a broad range of disciplines. Its rich bibliometric data and advanced search capabilities allow for precise and comprehensive literature searches. In addition, it indexes commonly cited sources in the field, including IEEE Xplore, Elsevier, and Springer, ensuring access to high-quality research.

2.3.1 Automatic retrieval procedure

Utilizing Scopus' Search Application Programming Interface (API), automated queries were generated to efficiently retrieve works of high relevance to the specific topic at hand. This was achieved using the Python library *pybliometrics* [101], which forms requests into a search matrix directed to the Scopus server, as explained in Section 2.3.2. The search was limited to the title, abstract, and keywords of the articles, ensuring not only avoidance of manual retrieval of pertinent works but also elimination of duplicate results. Following the initial retrieval, the results were manually filtered according to specific selection criteria outlined in Section 2.3.3. To further enrich the collection of references, both “retrospective” and “prospective” methods were employed. The “retrospective” method involved a manual inspection of references contained within the shortlisted publications. The “prospective” method, on the other hand, involved articles that cited the shortlisted publications if they met the selection criteria. The resulting bibliography comprises the articles that were retained following these comprehensive filtering stages.

2.3.2 Search matrix

The search matrix was built by combining keywords and enriching each keyword with corresponding alternatives. Five keywords were considered to cover the scope of interest for this literature study: ML algorithms, robot structure, system description, interventional type, and device type. The first keyword ensures that the related work uses ML algorithms. The second and third keywords limit the retrieved literature to the field of FSIRs. The last two keywords narrow the search to typical clinical applications and different types of FSIRs.

Table 2.3: Keywords for search. The asterisk wildcard symbol* represent any number of additional characters, allowing for the inclusion of all variations stemming from the base word.

Algorithm		*AND*	**Characteristics**		*AND*	**System**		*AND*	**Intervention**		OR	**Device**	
machine learning	*OR*		flex*	*OR*		robot*	*OR*		*vascular*	*OR*		catheter*	*OR*
deep learning	*OR*		continu*	*OR*		automat*	*OR*		percutaneous*	*OR*		*scop*	*OR*
reinforcement learning	*OR*		compliant*	*OR*		assist*	*OR*		cardiac*	*OR*		needle	*OR*
learning-based	*OR*		soft*	*OR*		autonomous	*OR*		colono*	*OR*		dissect*	*OR*
Hidden Markov	*OR*		snake*			guide*	*OR*		uretero*	*OR*		*wire	*OR*
SVM	*OR*								gastro*	*OR*		grasp*	*OR*
Bayesian	*OR*								broncho*	*OR*			
Long Short-Term Memory	*OR*												
Neural Network*	*OR*												
CNN	*OR*												
RNN	*OR*												
LSTM	*OR*												

The Scopus query was guided by the use of two logical operators, *AND* and *OR*. The asterisk wildcard symbol, which represents any number of additional characters, was used to include all variations of a base word. This efficient approach allowed for a thorough exploration of the extensive search space. It is believed that the adopted strategy encompasses most of the relevant literature on this topic. To effectively examine the search space, it is vital to incorporate as many keyword alternatives as possible. However, expanding alternatives may also yield more unrelated papers. Hence, a trade-off was made to limit the list of alternative keywords to the ones most relevant to the topic. The final search matrix, formed using refined keywords and logical operators, is illustrated in Table 2.3.

2.3.3 Selection criteria

In principle, the retrieved results should be closely aligned with the scope of the review paper. However, even if a paper matches the search matrix well, and is related to both ML and FSIRs, it could still be out of scope for this review. The papers are considered eligible only if the authors adopted ML algorithms directly or indirectly in perception, modeling, control, and navigation of the FSIRs, while the papers about pure imaging processing are considered outside the scope of this paper. *Case 1*: Publications on ML-based image processing [102, 103, 104] on the images fetched by a FSIRs are considered out of scope, because these are pure imaging processing papers. Information from imaging processing is not applied to perception, modeling, control, and navigation of FSIRs. *Case 2*: Some works use ML algorithms for visual servoing of FSIRs [78, 105, 106]. Although ML is also used in image processing, the outputs of the ML algorithm were eventually used to control FSIRs. Therefore, these papers are considered to be in scope.

2.4 Machine learning applications in FSIR

As introduced in Section 2.1.2, ML approaches are widely used in all four tasks/phases considered in this work to solve the current challenges of FSIRs (see Fig. 2.5a) from different perspectives. In the perception phase, ML algorithms could calibrate or model sensors [44], detect collision and estimate the contact force with the environment [107, 108, 109], reconstruct the shape of FSIRs [110], or localize the tip [111]. During the navigation phase, ML algorithms generate feasible paths e.g. based on RL [68], learn motion primitives [47], and optimize planned trajectories based on human demonstrations [54, 112]. In the control phase, ML algorithms can predict control commands based on the learned IK model [75, 113], refine the parameters of the IK-based controller or the PID controllers [66, 114], and derive control signals by trial-and-error [69, 115]. Refining the information from Fig. 2.4(B) and Fig. 2.5a, among all tasks, the control task has showcased the highest diversity of approaches, with a significant number of distinct algorithms observed. In terms of modeling, ML algorithms are used to model nonlinearities within the continuum FSIRs [85, 91, 116, 117], model kinematics/dynamics [64, 118, 119], or to predict distal end status based on proximal information [89, 92]. In certain scenarios, ML algorithms can be utilized across different stages to attain general control of the robotic system [66]. Figure 2.4 provides a comprehensive demonstration of the prevalence of various ML algorithms on different tasks.

Regardless of the task considered, since all ML methods can be seen as data-driven approaches, their performance is strongly influenced by the training data. It is therefore essential to pay attention to data collection, which can be carried out in different ways depending on the complexity of the tasks. In lower-level tasks, such as modeling and control, data are usually collected directly from proprioceptive

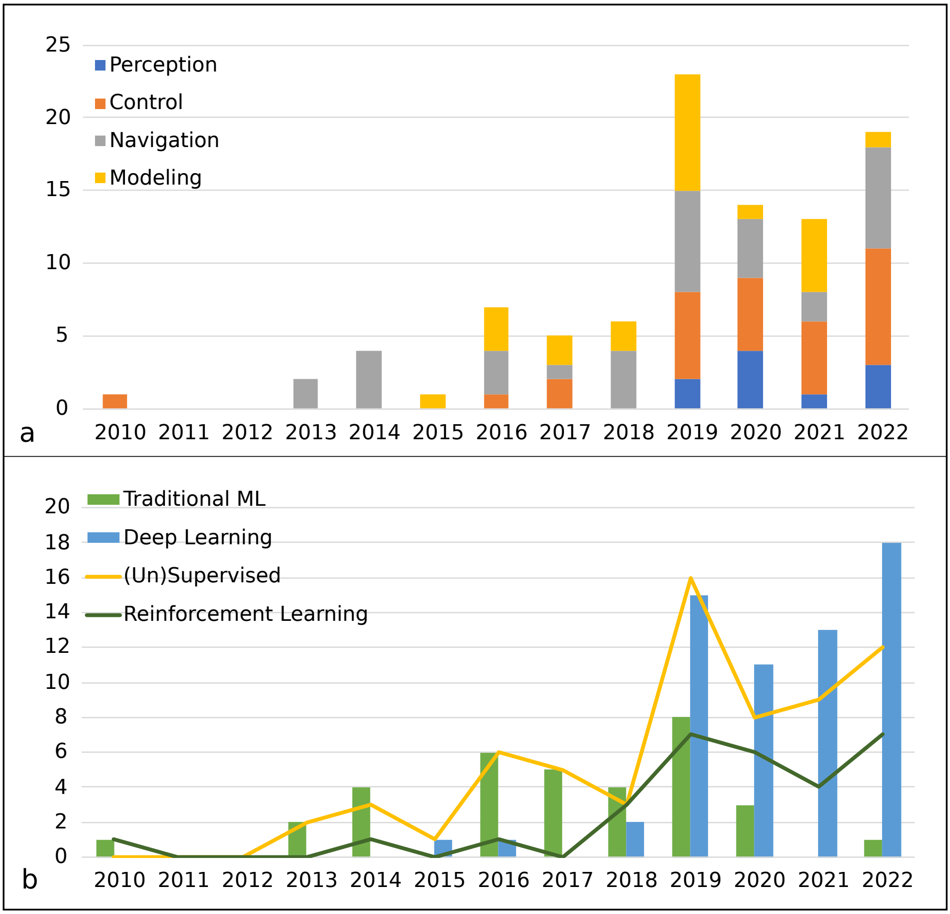


Figure 2.5: The trend in ML algorithm usage for various FSIRs tasks from 2010 to 2022 is illustrated over the years: (a) the number of publications of each task; (b) the count of publications in terms of the proposed 2D classification: Traditional ML, DL, supervised/unsupervised learning, and reinforcement learning.

sensors that are embedded in FSIRs [64, 75, 88] or placed in the environment [116, 120]. For more complex tasks, such as navigation, the fusion of proprioceptive and exteroceptive information [78, 121] plays an essential role in estimating the state of FSIRs while considering the interaction between FSIRs and the environment. In addition, demonstration data from human experts, animals [122] or even a rigid link robot [46, 123] can be used to teach FSIRs to complete high-level tasks.

If the configuration of FSIRs, such as the tip pose and position, shape, and contact state, could be determined, these data could be used to help guide navigation and control FSIRs in a more precise and delicate manner. Several methods have been proposed in the literature to improve the awareness of FSIRs during the procedure. According to Table 2.4, DL methods (MLP, CNN, and Encoder-decoder) are quite popular in the perception task. MLP is the most common ML approach, followed by CNN due to its strength in extracting information from images. Nevertheless, some traditional ML algorithms, e.g., gradient boosting and GPR, have also been employed in recent years. In this section, different ML-based pose/contact estimation methods are discussed.

Pose estimation

2.4.1 Perception

Cardiac catheterization is usually based on fluoroscopy guidance, but the generated ionizing radiation is harmful to both the patient and interventionalists. Another limitation of fluoroscopy is that it only provides 2D visualization, forcing interventionalists to correlate the obtained 2D information of FSIRs with the 3D patient's anatomy in their mind. Due to these disadvantages, different sensor modalities have been investigated to replace the traditional fluoroscopic system. Several ML-based approaches using the information provided by sensors have been developed to track the pose of FSIRs in real time (including position tracking and shape sensing) during MIPs.

Distributed sensors are commonly used to sense the shape of the FSIRs since they bypass the problem of line-of-sight (i.e., inside the patient's body). A framework for 3D shape reconstruction of a three-segment soft robot using DL with feedback from proprioceptive sensor skin is described in [124]. The sensor skin was made of conductive silicone using kirigami and a novel kinematic description linking the entire soft robot's structure to low-dimensional configuration parameters was also proposed. The study demonstrates the effectiveness of an LSTM network in learning the relationship between the sensor's resistance and the configuration parameters.

Fiber Bragg Grating (FBG) is increasingly being used in shape sensing of FSIRs due to its tiny size, biocompatibility, and safety (e.g., being free from the risk of electrocution). The traditional FBG-based shape sensing method starts by calculating the curvature from the measured wavelength shift of FBG sensors distributed along the fiber length. The calculated curvature is then integrated to reconstruct the 3D shape of the fiber. One of the major problems in traditional FBG-based shape sensing methods comes from the inaccuracy in estimating the curvature from the measured wavelength shifts of FBG sensors. Several characteristic parameters of the fiber, *e.g.* strain sensitive coefficient, angles between cores and distance between the central and outer cores, are required to calculate the curvature. However, these parameters could change during the assembly procedure, which requires a complex identification process. To overcome this problem, Manavi

Table 2.4: Applications of ML algorithm in the perception task (yellow for traditional ML, green for DL)

Subtask	Reference	Involved ML Method	Actuation Method	Type of Robot	DoF	Input	Evaluation task	Accuracy
Pose estimation	[124]	LSTM, MLP	Fluidic actuator	CR	2	Resitances of Kirigami Sensors	Estimate posture parameters of the robot	RMSE of posture parameters prediction: 1.27 mm (LSTM), 5.6 mm (MLP)
	[125]	MLP	/	Guidewire	/	Wavelength shifts from FBG sensors	Predict tip position	Predict the shape of a fiber sensor consisting of five FBG triplets with less than 6 mm tip error
	[44]	DL, LSTM	Cable-driven	CR	2	Wavelength shifts from FBG sensors	Estimate 3D shape of the CR	Mean error of 3D shape estimation: 0.66 mm (DNN), 0.69 mm (LSTM); Mean error of distal end position estimation: 0.45 mm (DNN), 0.48 mm (LSTM)
	[126]	DL	Fluidic actuator	CR	2	Wavelength shifts from FBG sensors	Estimate 3D shape of the CR	Mean error of 3D shape estimation: 0.24 mm (3D free-space experiment), 0.49 mm (2D constrained-space experiments)
Contact estimation	[127]	GPR	Fluidic actuator	CR	2	Robot's configuration and actuation	Detect contact state and estimate force direction	Contact state estimation: correctly detects the interactions any time they occur (100% success rate); force direction estimation: the average alignment of estimated and ground truth force direction is 0.95 (where 1 is perfectly aligned and 0 is orthogonal)
	[107]	Gradient Boosting	Cable-driven	CR	2	Wavelength shifts from FBG sensors	Detect contact state	Contact state estimation: successful detection of collision with hard and soft obstacles within less than 300 milliseconds.
	[128]	AutoEncoder	Cable-driven	CR	2	wavelength shifts from FBG sensors	Estimate contact stage and contact location	Estimate contact stage accuracy: 100% (in at most approximately 1.08 s); Mean contact localization error: 2.3 mm for a 70 mm long CR
	[129]	CNN	Cable-driven	Ablation catheter	2	Optical coherence tomography images	Estimate tip contact and orientation stage	Estimate tip contact stage accuracy: 99.96%; Estimate orientation state accuracy: 92.88%
	[108]	MLP, ELM	Fluidic actuator	CTR	2	Actuation input signal and current shape	Estimate tip contact force	RMSE of tip contact force estimation: approximately 65 mN

et al. proposed to train an ANN to extract shape information directly from the Edge-FBG spectrum [125]. The trained ANN can reconstruct the shape of a fiber sensor made up of five Edge-FBG triplets with an error of less than 6 mm at the tip. To further avoid the consequence of integrating inaccurate curvatures estimated by the traditional FBG-based shape sensing method, Sefati *et al.* [44] proposed an ANN-based method to directly estimate the tip pose of a continuum manipulator from the measured wavelength shifts of FBG sensors. The continuum manipulator is modeled as a pseudo-rigid body. Then a constrained optimization problem is formulated to solve the joint angles of the continuum manipulator to reconstruct the shape. The proposed ANN-based shape sensing method has improved shape sensing accuracy by approximately 58% compared to the traditional FBG-based shape sensing method. However, the approach suggested in [44] uses an optimization process that is computationally demanding, complicating its application to a continuum robot with multiple DOFs. Ha *et al.* presented an ANN model that can directly estimate curvatures at a discrete point along the length of the continuum robot providing the wavelength shifts measured by the FBG sensors in [126]. The experimental results show that the method presented in [126] outperforms the traditional FBG-based shape sensing method, which relies on the assumption of the geometry of the fiber core.

Magnetic field-based tip pose estimation methods have generated substantial interest in clinical applications due to their cost-effectiveness and capability to operate without requiring direct line-of-sight. A permanent magnet-based localization for growing robots was presented in [111]. The growing robot is a type of continuum robot that achieves locomotion by extending from its tip. The tracking system includes a permanent magnet integrated into the robot tip, and an array of magneto-inductive sensors that are placed underneath the robot to gauge the variation in the magnetic field when the robot navigates its workspace. In this work, a hybrid strategy was proposed that combines a dipole model-based localization with trained ANN. This strategy compensates for the measured flux density of each sensor in order to improve the accuracy of the tip pose estimation. The presented hybrid approach achieves a position error of 4.3 ± 2.3 mm in a 5-degree-of-freedom (DOF) setting. The hybrid model improves position tracking accuracy by approximately 37% (compared to the model-based localization method) when the magnet moves at high speed (*i.e.* 255 mm/s).

Contact estimation

During MIPs, knowledge about the interaction force between FSIRs and the surrounding environment is desirable, in addition to the posture of FSIRs. Information on current contact forces helps improve the awareness of clinicians or the autonomous system, which allows them to make better decisions during navigation. Incorporation of force sensors on FSIRs to measure either tip force, distributed force, or contact state is difficult considering the size limitations and constraints related to cleaning and sterilization. A model-based component was presented as an alternative to direct sensing in [127]. The nonlinear disturbance observer, with corrections learned from data via the ML method, namely Gaussian Process Regressions, can estimate the contact state and force direction. The method in [127] only requires knowledge of the robot's posture and actuation signals. To measure its efficiency, the normalized scalar product of the estimated force direction with the ground truth was evaluated. The average alignment yielded the result of approximately 0.92 (where 1 is perfectly aligned) with a standard deviation of 0.08. Another approach was proposed in [107] to estimate the contact state of a continuum manipulator. This work proposed a trained ANN to directly detect collision from the measured wavelength shifts of the integrated FBG sensors. The experimental results demonstrated successful detection of collisions with hard and soft obstacles in less than 300 milliseconds. A data-driven approach to estimate not only the contact state, but also the contact location, is proposed in [128]. The proposed method requires only measurement of curvatures along the length of the robot. In this work, the curvatures are measured by a multicore optical fiber embedded in the central channel. The experimental results show a mean average localization error of 2.3 mm for a flexible robot of 70 mm long. Knowledge of the interaction force is not only useful for safe navigation, but also important for other tasks that require maintaining contact between FSIRs and tissue for a period of time, such as catheter ablation treatment. Yu *et al.* proposed a method to estimate contact and orientation based on direct endomyocardial imaging acquired by optical coherence tomography [129]. A CNN-based two-stage classifier was developed to provide an intra-procedure assessment of contact parameters, achieving precision of 99.96% and 92.88% for the estimation of contact and orientation, respectively. Unlike [107, 127, 129] which mainly concentrates on estimating the contact state, a virtual sensor was developed to estimate the contact force at the tip of concentric tube continuum robots through supervised learning [108]. This work also shows the usefulness of the transfer learning approach based on deep direct cascade learning. The deep direct cascade network was pre-trained with synthetic data generated by a simulated model of the robot before training with real data. The accuracy in the estimation of the contact force of the tip was 8.471 ± 1.389 mN [108].

2.4.2 Modeling

In the realm of controlling conventional rigid link robots, researchers have developed a variety of analytical models. These models effectively describe the kinematics and dynamics of the robot, providing a detailed understanding of their operation. However, when we change focus to FSIRs, the scenario changes considerably. Unlike rigid link robots, analytical models for FSIRs are inherently more complicated and, to some extent, less precise. This is mainly due to the intrinsically non-linear nature of FSIRs. Furthermore, FSIRs can potentially possess a higher number of degrees of freedom (DOF), which subsequently increases the complexity of these analytical models. Thus, comprehensive modeling of FSIRs represents a challenging facet in robotics research.

The topics discussed in this section are related to the modeling of the characteristics of FSIRs. These characteristics could include kinematics, dynamics, or shape/pose, as outlined in Table 2.5. Efficient models of characteristics can enhance our comprehension of the FSIRs, thereby refining control accuracy/precision and environmental adaptability.

According to Fig. 2.4(B), MLP is the most widely used algorithm that accounts for approximately 29% of all the ML algorithms used in the modeling task due to its strong capability of modeling nonlinearities and convenience for implementation. Other DL algorithms such as LSTM and CNN are also frequently utilized.

Model kinematics/dynamics

ML methods are used to represent FSIRs's kinematics, bypassing intricate analyses of robot attributes and formulas. These techniques can be easily adapted across various FSIRs types, offering a versatile modeling strategy. In addition, the ML approach, as a data-driven method, collects training data from the real robot. This procedure identifies patterns of the robotic system in which nonlinear phenomena are also incorporated. This is more useful, compared to traditional kinematics/dynamics models, for continuum robots, in which the nonlinearities are noticeable. Note that in this subsection, the included articles discuss the modeling of the robot kinematics or dynamics when nonlinearities are not a major challenge in those FSIRs such as multi-joint surgical robots [119], or CTRs [63], or when nonlinearities compensation was not claimed as a contribution by the authors.

Conventional ML methods, such as density models and even regression, have been explored to address these issues. A comparison between GMR and k-NN regression is presented in [76], to learn the inverse kinematics (IK) of a Tendon-Sheath-Mechanism (TSM) robot. The training data consists of pairwise motor movements and end-effector states generated by human demonstration. The average prediction accuracy (considering both actuators used in the TSM robot) is 91.2% and 93.2% for GMR and k-NN, respectively. A subsequent study conducted by [64] included Extreme Learning Machine (ELM) for comparison. The RMS errors of GMR, k-NN regression, and ELM on trajectory following experiments are 2.5556 mm, 2.1527 mm, and 2.3277 mm, respectively.

In [113], another regression method is explored with the objective of predicting the control signal. Here, the focus is on directly approximating the Inverse Kinematics (IK) for a redundantly actuated, fluidics-driven soft robot. The research introduces a comprehensive global IK model, constructed as a strategic weighted mixture of numerous local inverse models. This methodology leverages the benefits of Locally Weighted Projection Regression (LWPR), which operates in an online mode. With

Table 2.5: Applications of ML algorithms in the modeling task (yellow for traditional ML, green for DL and purple for RL)

Subtask	Reference	Involved ML Method	Actuation Method	Type of Robot	DoF	Modeled Object	Evaluation task	Performance
Model Kinematics/Dynamics	[86]	MLP	Simulation	Continuum Robot (CR)	3	IK	Trajector following	RMSE: ca. 0.18 m and ca. 0.21 m (with last 3 joints fixed)
	[130]	MLP	Simulation	Concentric Tube	5	FK, IK	Predicting the tip position, tube extension and orientation	Mean tip error: 0.2 mm, 0.002 difference (FK); Error in tube extension: 0.8 mm, errors in tube rotation 0.1° (IK)
	[63]	MLP	Cable-driven	Concentric Tube	6	FK, IK	Predicting the tip position, tube translation and orientation	RMSE: 2.3 mm in position, 1.1° in orientation (forward) 4.0 mm in position, 8.3° in orientation (inverse)
	[131]	MLP	Cable-driven	Multi-joint	7	FK, IK	Predicting the trajectory	RMSE: 3.047e-07 (forward) mean absolute error: 1.8e-03 (inverse)
	[119]	LSTM	Cable-driven	Multi-joint	7	IK	2D trajectories following	Mean tracking error norm: 3.47 mm (Square)
	[118]	BNN	Cable-driven	Multi-joint	7	Kinematics, Dynamics	Predict tip position and motor torques	Test RMSE: [0.109, 0.145, 0.108]mm (kinematics); [0.037, 0.050, 0.050, 0.024, 0.025]Nm (dynamics)
	[76] [64]	GMM, KNNR	Cable-driven	Serpentine manipulator	2	IK	2D trajectories following	RMSE: 2.5556 mm (GMR), 2.1527 mm (KNNR)
	[113]	LWPR	Fluid-driven	CR	2	IK	3D trajectory following	Mean accuracy: ±2.21° and maximum ±7.49°; ±2.49° and maximum ±11.03°) with external disturbance
	[75]	GMM	Cable-driven	CR	3	FK, IK	3D trajectory following	RMSE and standard deviation: 1.379 ±0.637 mm (forward), RMSE 1.909 ±1.067 mm (inverse)
Model Nonlinearities and Hysteresis	[85]	MLP	Cable-driven	CR	2	Hysteresis	Hysteresis modeling	Error distribution of 0.11 with 81.8% variability in hysteresis
	[91]	LSTM	Fluid-driven	CR	1	Hysteresis	Predict the bending angle with hysteresis	RMSE: 0.36°; MAE: 1.23°; NRMSE: 1.17%
	[117]	LSTM, ELM	Cable-driven	CR	2	Backlash	3D trajector following	RMSE: [9.6, 9.1, 9.34]° (LSTM); [12.9, 12.6, 12.0]° (ELM)
	[132]	CNN	Cable-driven	Multi-joint	2	Hysteresis	Pose estimation with hysteresis	Average hysteresis reduction rate of RMSE: 60.24±0.37% at θ_1 and 65.15±0.65% at θ_2
	[133]	MLP	Cable-driven	Multi-joint	2	Hysteresis	Estimate the joint angle of the robot	Maximum hysteresis sizes for 4 degrees: [4.39±1.0, 7.3±0.9, 8.19±0.9, 12.4±1.5]°
	[116]	ELM	Cable-driven	CR	3	IK Hysteresis	2D and 3D trajectories following	RMSE: 0.55 mm (2D), 2.03 mm (3D)
	[134]	ANFIS	Cable-driven	Adapted commercialized catheter	2	Backlash	Path following	Mean displacement: 1.997 mm ±0.849 mm; Mean instantaneous velocity: 3.669×10^{-4} mm/sec ±10.767 mm/sec
	[135]	PoWER	Cable-driven	CR	3	IK, nonlinearity	2D trajectories following	RMSE after RL refinement: 1.101 mm
	[66]	PoWER, Liner Regression	Cable-driven	Serpentine Manipulator	3	IK, nonlinearity	Tube insertion and circle following	Returns of rollout: 0.986 (tube insertion) and 0.981 (circle following)
Model Distal-end force Feedback	[120]	CNN	/	Guidewire	2	Force mode	Obstacle avoiding	Maximum operating forces: 0.24 to 0.79 N
	[92]	LSTM, MLP	Cable-driven	CR	2	Force	Predict force hysteresis	Average RMSE: 0.1364 N
	[89]	CNN, LSTM	Cable-driven	CR	2	Force	Predict force hysteresis	Average RMSE: 0.1711 N

this approach, the model is not only robust but also adaptable, capable of real-time updates, offering a novel and efficient perspective on control strategies in soft robotics. The proposed approach achieved a mean tracking error of ±1.79° based only on the prepared simulation data. The performance eventually improves to ±0.90° with online learning. In [75], the Dynamic Gaussian Mixture Model [136], a modified version of GMM, is employed to represent the forward kinematics (FK) and IK of a robotic catheter using its present state, next state, and related control commands.

ANN also demonstrates its capability to model the kinematics of FSIRs. Especially when FSIRs has higher number of DOFs, DL is more widely used than traditional ML. This trend can be seen in Table 2.5. In non-redundant continuum robots for nonclinical applications, previous work [137] has demonstrated the capability of ANN in modeling kinematics and showed superior performance compared to conventional methods. MLP is most employed in the modeling task for FSIRs, and it is also the most popular algorithm across the reviewed articles as shown by Fig. 2.4(A). A related work appeared for CTR [130]. Its FK and IK were learned using an MLP. However, this work was only validated in simulation. A subsequent work [63] proposed MLP-based learning of the FK and IK for a 6-DOF concentric tube continuum robot. Unlike [130], the data in [63] were obtained in a real-world robotic platform. Thuruthe *et al.* addressed the challenge of IK learning for continuum/soft robots by optimizing the training data representation [86]. They employed MLP as a learning technique using data from a simulated kinematic model. Additionally, they introduced a trigonometric joint description for enhanced feature representation in the learning process. The learned model of FK gave a tip error of less than 2.3 mm in position and 1.1° in orientation. As for IK, the model achieved a translation actuator error of 4.0 mm and a rotation actuator error of 8.3°. The author claimed that the proposed feature description improved the prediction accuracy of the learned model. Li *et al.* used three types of ANN with different loss functions to model the kinematics of a multi-joint tendon-driven flexible surgical robot called Micro-IGES [131]. This work was extended by using an LSTM to model the kinematics of the Micro-IGES robot [119]. The LSTM network is capable of learning the relationship between the actuation space and the task space. The results showed that the robot could achieve better control accuracy with the LSTM than with Denavit–Hartenberg models. In Micro-IGES, a Bayesian Neural Network (BNN) was also used to learn the kinematics and dynamics of the robot [118]. In addition to modeling, the BNN is also able to provide information about the epistemic uncertainties of the learned system. Compared to MLP, BNN is less prone to overfitting [118]. Real-world experiments revealed that the tip positioning accuracy with BNN outperforms that with a pseudo-inverse kinematic controller. However, the BNN needs to be compared to other state-of-the-art models, e.g., MLP and LSTM in order to fully understand its modeling capabilities.

Model nonlinearities and hysteresis

Nonlinearities, such as hysteresis, deadzone, and backlash, pose significant challenges in the control of glsfsir. These complexities have caught the keen interest of researchers within this field. It should be acknowledged that a multitude of articles in this section essentially focus on modeling the Inverse Kinematics (IK). However, the distinguishing factor of the articles in this section, in comparison to the previous, lies in the degree of non-linearity of the FSIRs under discussion. Here, the non-linearity of FSIRs is deemed more pronounced, and is thus treated as a major challenge by the authors. Therefore, while maintaining the learning of IK, this section underscores the significant role of nonlinearities within the broader context of FSIRs research. Nonlinearities in FSIRs arise from elements such as their flexibility, interactions between segments, and internal friction. Such nonlinearities lead to differences between the anticipated motions at the distal end and the actual actions at the distal segment. This hinders the effective modeling and control of FSIRs during MIPs. To learn such non-linear relations, the training data is often collected from the robot movements, which reflect non-linear behaviors of the robot.

In [116], an ELM-based offline learning method was introduced for hysteresis modeling. Separate ELMs were used to fit models, which collectively determined the joint positions. This method was validated on trajectory following experiments, and achieved an RMS error of 2.03 mm for 3D trajectories. Rather than modeling the causes of nonlinearity separately, Omisore *et al.* proposed to model the nonlinear relationship using a single MLP with relevant variables, e.g., displacement and velocity of the proximal part [85]. The last output was used again as input. In their ANN structure, a naive RNN concept is implemented, namely they take the last output as input. In their approach, they consider only the most recent state for time dependence. Wu *et al.* employed LSTM to model rate-dependent hysteresis in a robotic catheter driven by pneumatic actuation [91]. Training data consists of bending angle of the catheter tip (output) and pressure (input), which are collected to train the LSTM. The performance of LSTM that has been trained is compared to that of a state-of-the-art analytical model known as the Deadband Rate-Dependent Prandtl-Ishlinskii (DRDPI) model and the traditional ML method SVR in a trajectory following experiment. The results showed that the LSTM outperformed the analytical model and the SVR model by 60.1% and 36.0%, respectively, in an arbitrarily varying trajectory. While the aforementioned study provides insight into the modeling of hysteresis, it does not extend to its compensation. This unresolved challenge is addressed in a subsequent work [138]. This work designs an open-loop controller based on an LSTM to control and compensate for hysteresis in a catheter system. The proposed control-LSTM is used as a feedforward free-space catheter controller robust to severe hysteresis. It is precise and has a simple training procedure. Furthermore, the control-LSTM does not require a separate inversion step when employed as a controller as would be needed by most traditional analytical hysteresis models such as [139]. Similarly, the backlash of a two-DoF steerable endoscope was modeled using LSTM in [117]. The LSTM takes bending angles in 2 DOFs as input and predicts the 4-way backlash of the endoscope.

Omisore *et al.* initially examined the various factors impacting the backlash gap, identifying the most influential ones [134]. They then developed a closed-loop control system to compensate for backlash. This compensation was made possible through an Adaptive Neuro Fuzzy Inference System (ANFIS) [140] model, which predicts potential backlash, and a force modulation model that monitors contact forces between the catheter tip and vasculature. In their in vitro experiment, the mean absolute deviations of the input signal, and outputs of displacement and instantaneous velocities, are 0.85 mm and 10.77 mm/sec, respectively.

Visual feedback could also be used to compensate for the hysteresis. A hysteresis compensator with learning-based pose estimation is proposed in [132]. Tip pose estimation was implemented based on a siamese CNN to reduce hysteresis by adjusting the position command with a supplemental movement, similar to the instinctive adjustments of clinicians with their visual feedback. Validation on the testbed showed that the method can reduce hysteresis by up to 71.4%. The visual method mentioned above was combined with a kinematic-based joint angle estimation method to further compensate for the hysteresis in the same robot [133]. Hysteresis could be reduced to less than 5° when the sheath configuration is 0° or 90°.

Another approach that combines an analytical model with RL to solve the problem of nonlinearity is presented in [66] and [135], where an RL algorithm is employed to compensate for a 2-DOF TSM's nonlinearities. Chen *et al.* proposed a three-layered approach that combines motion planning, solver of IK, and refinement of RL [66]. The initial layer involves the recording and modeling of a human demonstration of controlling the TSM using GMMs. In the second layer, the TSM performs the commands derived by the GMR. The GMR is derived from the GMMs, as explained in Section 2.2.2. In addition, the positions of the end effector are recorded for learning the IK model with linear regression. It is worth noting that the GMM-GMR approach adopted here is the most popular among all the traditional ML methods reviewed in this article according to Fig. 2.4(A). In the third layer, RL (Policy Learning by Weighting Exploration with the Returns, PoWER) is employed to refine the IK model by learning from the tracking trajectories generated by TSM. However, the refined IK model was only validated in simulation by a circle drawing experiment. This work is extended to [135] by conducting an evaluation in real conditions. A serpentine manipulator is controlled to follow different trajectories using the refined IK model. The RMS errors produced by the IK model before and after the RL refinement are 9.639 mm and 1.101 mm, respectively.

Model distal-end force feedback

Recognizing the forces applied on the distal end of the FSIRs is essential to obtain a correct estimation of the current robot state. This information can be utilized in robot control to ensure safety in MIPs [120]. However, obtaining precise force feedback is difficult due to challenges in placing sensors at the distal tip and the intrinsic nonlinearities of FSIRs. Consequently, there is growing interest among researchers in sensorless methods for estimating distal-tip force, which can also aid in creating haptic feedback.

Li *et al.* explored DL techniques to estimate the TSMs' distal force using only measurements from the proximal end [92]. They compared LSTM and MLP against traditional model-based methods through a tissue manipulation task. Additionally, they introduced a two-phase data-driven approach, rooted in ANN, to forecast the dynamic force at the distal tip of a flexible robot [89]. In the initial phase, both probing signal data and proximal-end force feedback are transformed into 2D visuals. A CNN classifies and estimates the sheath's bending angle from these images. Following this classification, the second phase employs the two top-performing LSTM models from each class to dynamically gauge the robot's distal end force.

2.4.3 Control

Managing control is a primary challenge for any robotic systme. The flexible nature of FSIRs introduces higher DOFs and reduced rigidity, which complicates its comprehension and control. Developing a control schema that ensures FSIRs applicability to dynamic environments with high control accuracy is essential for surgical safety. This section summarizes three control manners: feedforward control using models, feedback control and RL-based control, as shown in the Table 2.6. The feedforward control with models eliminates sensors mounted on the tip of the FSIRs, which is considered challenging due to the size limitation and sterilization constraints. The feedforward control can work in some free-space scenarios, such as inside the heart. However, the control commands adjustment is not error-based, which makes it difficult to be used in a constrained environment. Feedback control relies on feedback from sensors based on various physical principles, e.g., visual, electromagnetic, or optical. RL is also used to train controllers, which translates the high-level commands to motor skills from the interaction between the robot and the environment.

According to Fig. 2.4(a), LSTM is the most popular algorithm in control. Although LSTM enjoys popularity, it does not dominate the task. The combined occurrences of the first four algorithms (LSTM, CNN, DDPG and GMM-GMR) in the control task account for nearly half of the total. The control task has the largest variety of algorithms among tasks and there is not a mainstream algorithm in the control task.

Feed-forward control using models in Section 2.4.2

As discussed in the "modeling" part, ML algorithms played a crucial role in learning the IK model of a FSIRs. The learned IK model accepts a desired tip trajectory as input, and predicts control commands for the actuators. With the learned IK model, a feed-forward controller could be implemented with minor effort.

A few articles discussed in the previous subsection of modeling kinematics use this feed-forward control to validate their model on FSIRs [31, 64, 76, 86, 138]. These works integrated their learned IK models in the robot feed-forward control and validated the performance by trajectory following experiments. Most validation experiments are conducted in free-space [64, 76, 116, 118, 119, 131, 135] because the learned IK model cannot consider perturbations from the environmental. Nevertheless, compared with conventional feed-forward control based on analytical models, this data-based approach is easier to implement. The ML approaches only require data to be collected on the FSIRs without a deep understanding of the system. The collected data often contains a lot of learnable patterns and information. The ML can learn all these factors in one shot. On the contrary, the analytic models require deep knowledge of the system. This increases the workload if the structure/configuration of the FSIRs is changed. In addition, some factors are difficult to be observed so that they are not considered by the analytic model. However, direct elimination of these factors largely influences the control precision.

Feedback control

Analytical models or learning-based models for the feedforward control of a robotic system could be accurate enough in open space. However, when applied to robots in the world with external disturbance and interacting with the environment, the error increases. Thus, it is not easy to achieve precise control on FSIRs solely relying on feedforward control. Feedback information on the current status of the robot is important to adjust the control commands for more accurate control.

Visual feedback Visual feedback from cameras contains instructive spatial information for FSIRs control. Cameras could be embedded on the distal end of robots, or placed in the environment to observe the procedure. However, as mentioned in Section 2.4.1, external cameras suffer from the line-of-sight problem and are therefore difficult to use in a clinical setting. From the visual feedback, the controller has to estimate the gap between current status and expected status, and decide how to adjust the actuators to reduce the gap.

Table 2.6: Applications of ML algorithms in the control task (yellow for traditional ML, green for DL and purple for RL)

Subtask	Reference	Involved ML Method	Actuation Method	Type of Robot	DoF	ML Role	Real-world Experiment	Simulated Experiment	Validation task	Results
Feed-forward Control using models	[118]	BNN	Cable-driven	multi-joint	7	Learn the IK to predict control signals	✓	✓	Predict tip position and motor torques	Test RMSE: [0.109, 0.145, 0.108]mm (kinematics); [0.037, 0.046, 0.050, 0.023, 0.022]Nm(dynamics)
	[131]	MLP	Cable-driven	multi-joint	7	Learn the IK to predict control signals	✓	x	Predict the trajectory	RMSE: 3.047e-07
	[119]	LSTM	Cable-driven	multi-joint	7	Learn the IK to predict control signals	✓	x	2D trajectories following	Mean tracking error norm: 3.47 mm (square trajectories)
	[76] [64]	GMM, KNNR, ELM	Cable-driven	CR	3	Learn the IK to predict control signals	✓	x	2D trajectories following	RMSE: 2.56 mm (GMR), 2.15 mm (KNNR), 2.33 mm (ELM)
	[116]	ELM	Cable-driven	CR	3	Learn the IK with hysteresis to predict control signals	✓	x	2D 3D trajectories following	RMSE: 0.55 mm for 2D, 2.03 mm for 3D
	[135]	PoWER	Cable-driven	CR	3	Learn the IK with nonlinearity to predict control signals	✓	✓	2D trajectories following	RMSE after RL refinement: 1.101 mm
Feedback Control	[105], [106]	RNN	Cable-driven	multi-joint	4[105], 6[106]	Quadratic programming solver for the kinematics-based control	× [105], ✓[106]	✓	Visual servoing	Error on 3 axes: < 0.2 mm; average computing time < 0.01 s [105]; Demonstrated the visual servo control [106]
	[88]	CNN	Fluidics-driven	CR	5	Predict the relative orientation of the placental surface for tip position control	✓	x	Distal tip alignment	RMSE: 5.93°
	[120]	CNN	/	Guidewire	2	Recognize the operating force mode for force control	✓	x	Obstacle avoiding	Maximum operating forces: 0.24 to 0.79 N
	[141]	CNN	/	/	/	Predict risk probability from force state and torque state	✓	x	Guidewire cannulation	Average operating force and average operating torque reduced by 20.80% and 14.20%
	[113]	LWPR	Fluidics-driven	CR	2	Learn the IK to predict control signals	✓	x	3D trajector following	Mean accuracy: ±2.21° and maximum ±7.49°; ±2.49° and maximum ±11.03° with external disturbance
	[75]	GMM	Cable-driven	Commercial catheter	3	Learn the IK to predict control signals	✓	✓	3D trajector following	RMSE and standard deviation: 1.379 ±0.637 mm in reality; 1.909 ±1.067 mm in simulation
	[77]	SVM, MLP*	Cable-driven	Tendon-sheath Mechanism	2	Learn the IK to predict control signals	✓	x	2D trajectories following	Average RMS error: 0.49 ±0.32 and 0.62 ±0.36 mm for the slow and fast speeds
Control with Reinforcement Learning	[142]	SARSA	DC motors	multi-joint	1	Learn to generate a policy to move the robot in the colon	✓	×	In-vivo colon endoscope	RL gave significantly better results in more than 50% colons compared to fixed input
	[115]	MA fuzzy Q-learning	Cable-driven	Concentric tube	2, 3**	Learn optimal policy of each controller	×	✓	Trajectory tracking	Correlations between achieved and desired trajectories in X and Y directions: 95% and 97% (2DOF); 94% and 97% (3DOF)
	[69]	DQN	Cable-driven	Commercial catheter	3	Learn to generate policy to control the catheter	✓	✓	Reaching targets	Average distance between catheter tip and target: 4.70±1.59 mm
	[98]	DDPG	Cable-driven	Concentric tube	4, 6, 8**	Learn the control policy for concentric tube robots	×	✓	Trajectory following	Mean error ranging from 0.31 to 4.35 mm
	[143]	DDPG	/	Concentric tube	6	Learn the control policy for concentric tube robots	×	✓	Trajectory following	Mean Cartesian error: 1.29 mm mean Cartesian error in the IK evaluation; mean tracking error: 1.37 mm (with noise)
	[114]	DQN	Cable-driven	Commercial catheter	4	Learn the control policy for concentric tube robots	×	✓	Point tracking	RMSE: 0.003 ±0.0058 mm

* Both SVM and MLP play essential roles in [77]. For the convenience of illustration, in this table, cells of [77] are filled with yellow (traditional ML).

** Experiments are conducted on robots with different DOFs.

[105, 106] proposed to use RNN in their visual servoing control scheme. The suggested method employs quadratic programming to navigate a flexible endoscope, integrating both kinematic and physical limitations. The RNN is used as a solver with finite convergence time to deliver a kinematic model suitable for closed-loop control. The effectiveness of the visual servoing control system based on accelerated RNN was tested in two simulations of distinct robotic platforms. The results showed that errors are less than 0.2 mm, and the average computing time for each time step is less than 0.01 seconds.

In addition to the promise of RNN-driven visual servoing, the profound ability of CNNs to process image data holds potential for enhancing FSIRs control. Deep learning techniques were introduced in [87, 144] to determine the position of a camera at the end of interventional instruments. These methods showcased potential uses in FSIRs control, a feat later realized in robotic fetoscopy research by [88]. In this study, a specially trained CNN predicted the relative angle of the placental surface based on fetoscopy images. This CNN was then incorporated into a collaborative control system, balancing the automated movement of the flexible tip with the broader motion directed by the operator. The proposed automatic tip control achieved a root-mean-square error of 5.93°. In [120], a CNN trained with surgical images is employed to estimate the likelihood of actions for control. For enhancing the safety of MIPs, a 1D CNN was developed using operational forces to identify the force mode. An alternative method for recognizing the force state in closed-loop control of a robotic catheter is presented in [141]. The authors used identifiers based on CNN to estimate the probability of irregular states for both force and torque, which were then used to control the catheter. The proposed method reduced the average operating force and torque by 20.80% and 14.20%, respectively. Due to the remarkable capability of CNN in effectively handling different data types, including images and force/torque measurements, CNN is also a popular algorithm and appeared in 12% of the reviewed articles as represented in Fig. 2.4(a).

The traditional ML methods are also used for visual feedback control in [75]. Closed-loop positional control is introduced based on a novel IK model integrating visual guidance. This IK model is learned by the dynamic Gaussian mixture model. The validation was carried out by following two 3D square trajectories. The mean RMSE and standard deviation of their trajectory following test is 1.379 ±0.637 mm.

Electromagnetic feedback Electromagnetic (EM) sensors are widely used in FSIRs because their small sizes ease the integration in the FSIRs tip. Electromagnetic feedback provides directly 3D spatial information, thus it is often used to measure the position and orientation of the tip of the FSIRs [77, 138].

Various ML methods are integrated at different stages of the control framework for FSIRs. In [77], Jolaei *et al.* developed a kinematic control scheme for soft tendon-driven catheters using a learning approach. They employed an SVM algorithm to determine which is the activated tendon. The SVM classification results are used by four deep ANNs to determine the desired length of each tendon. In their trajectories following experiments, the average RMS error of the system is 0.49±0.32 mm and 0.62±0.36 mm for the slow and fast speeds, respectively.

Control with reinforcement learning

RL is frequently used to help robotic systems determine the best control strategies when dealing with intricate environments. The realm of FSIRs control has seen sustained interest in RL among researchers. As RL has evolved, the algorithms applied to FSIRs have transitioned from basic ones like Q-learning and SARSA (state–action–reward–state–action) to more advanced, ANN-augmented algorithms like DQN and DDPG.

An initial study by [142] employed Q-learning and SARSA to adapt the motor's input voltage in order to decide the direction and speed of a colon endoscope robot. The robot earns a positive reward for forward movement, and punishments are introduced to prevent undesirable situations, such as when the torque surges to risky levels, potentially twisting the colon. Both techniques were tested on six in-vitro colon models. Both algorithms gave better results in over 50% of the colons than a usually used constant velocity controller.

While robotic control via RL is typically approached as a singular agent interacting with an environment, the control of tendon-driven manipulators is conceptualized from a multi-agent reinforcement learning perspective in [115]. In their study, each DOF is viewed as an individual RL agent. They employed a multi-agent fuzzy Q-learning method capable of establishing the relationship between the manipulator's tip positions and the desired path. In their simulation-based trajectory following experiments, correlations between achieved and desired trajectories are above 94% considering both 2 DOFs and 3 DOFs manipulators.

As described in Section 2.4.1 and Section 2.4.2, ANN has proved to be one of the most used approaches in several FSIRs tasks, for example in perception and modeling. Similarly, the advances on DRL, benefited from the evolution of ANN, broadened the applicable scenarios of RL in the FSIRs context, since these algorithms are suitable for learning a complex MIP task using only the clinical data commonly available during a procedure. Conversely, the computational resources needed to train DRL agents have grown significantly, creating a hindrance to further applications.

In a simulation study conducted by You *et al.* as detailed in [69], a dueling DQN was employed to learn the control of a cardiac ablation catheter tip. The training of the agent was based on both angiographic images and position data. To bridge the disparity between the simulated and actual practice, random disturbances were introduced during the training process, enhancing control precision under real-world conditions. The RL agent achieved an average error of 4.70±1.59 mm between the catheter tip and the target. In [98], Iyengar *et al.* delved into the influence of extra noise in RL-driven CTR control. They assessed three different noise types. Using the DDPG algorithm, they evaluated its efficacy in formulating the optimal control strategy across three CTR configurations: 2-tube, 3-tube, and 4-tube. This RL-based IK controller achieved an average extension error of 0.44 mm and a rotation error of 0.3°. In their follow-up study, the authors of [143] also used DDPG to build a model-free IK solver for concentric robot control. Compared with [98], [143] worked with a longer concentric tube with a more complex model in simulation. Also, a goal-based curriculum function is utilized to decrease the training time. The proposed IK solver achieved a 1.29 mm mean Cartesian error in the IK evaluation, and a 1.37 mm mean tracking error in a noise-induced simulation.

In [114], a DRL-based approach is applied to adapt the PID control gains when the robotic catheter interacts with the environment. A DQN augmented by LSTM is used to learn the tuning policies. The temporal learning approach is used in this process to increase the sampling efficiency and update the target network of the DQN between episodes. The suggested method was tested in a simulated experiment to gauge axial movements. The result is an RMS error of 0.003 ±0.0058 mm, which outperforms the other methods (i.e., Ziegler-Nichols system and adaptive fuzzy tuning).

2.4.4 Navigation

Navigation is one of the commonly occurring medical phases where the interventional tool advances through body lumens or vessels to reach deeper sections of anatomy [40]. Manual navigation usually takes up much of the clinicians' concentration during the procedure while deploying the device. Operating deployed instruments is even more critical. FSIRs with a higher level of autonomy in navigation can save clinicians' energy from focusing on meticulous operations during the navigation phase. This allows clinicians to focus more on crucial tasks after reaching the intervention site. Nevertheless, autonomous navigation for FSIRs is challenging because the environments are deformable and dynamic.

As a high-level task, navigation can benefit from efficient perception, reliable modeling, and precise control. Therefore, some research work on navigation has partially discussed these topics. In this subsection, we mainly focus on one of the most common navigation tasks, namely motion planning. Research in various robotic fields suggests that initiating motion planning strategies is a foundational step towards autonomous navigation [156]. Motion planning refers to finding feasible trajectories or actions traversing the area between a starting state and a goal state while bypassing obstacles and avoiding unwanted zones. Regarding MIPs involving FSIRs, the kinematic and geometric restrictions of FSIRs also impose constraints on the planning.

Table 2.7: Applications of ML techniques in the navigation task (yellow for traditional ML, green for DL and purple for RL)

Method	Reference	Specific Algorithm	Type of Interventions	Model Input	Model Output	Local Motion Planning	Trajectory Optimization	Learning Motion Primitives	Kinematics Constraints	Real-world Experiment	Simulated Experiment
LfD	[55]	GMM(EM)-GMR	Endovascular catheterization	Proximal data from expert motion	Motions	×	×	✓	×	✓	×
	[53]	HMM(EM)	Endovascular catheterization	Proximal data from expert motion	Motions	×	×	✓	✓	✓	×
	[112]	DMPs	Endovascular catheterization	Proximal and distal data from expert motion	Trajectories	x	✓	✓	✓	✓	x
	[47]	GMM(EM)-GMR	Endovascular catheterization	Proximal and distal data from expert motion	Trajectories	×	✓	✓	×	✓	×
	[52]	GAIL	Endovascular catheterization	Catheter states, manipulator motions	Motions	✓	×	✓	×	✓	×
	[145]	GAIL	Neurosurgery	Needle states	Trajectories	✓	✓	✓	✓	×	✓
	[51]	GMM-GMR	Laparoscopy	Tip trajectories from expert	Trajectories	×	✓	✓	✓	×	✓
	[146]	DP-Means	Laparoscopy	Tip trajectories from expert	Trajectories	×	✓	✓	✓	×	✓
	[45]	GMM-GMR	MIS	Demonstrated trajectories	Executable paths	×	✓	×	✓	✓	×
	[147]	GMM-GMR	MIS	Demonstrated trajectories	Motor trajectories	×	×	✓	×	✓	×
	[74]	GMM-GMR	Keyhole surgery	Demonstrated trajectories	Executable paths	×	✓	×	✓	✓	×
	[123], [46]	GMM-GMR	Laparoscopy	Contexts/phases of the task	Motions	×	×	✓	✓	×	✓
	[148], [123]	GMM-GMR	Laparoscopy	Octopus movements	Motions	×	×	✓	✓	×	✓
RL	[149]	Q-learning	Endovascular catheterization	2D mesh	Motions	✓	×	×	×	×	✓
	[99]	DQN, DDPG	Endovascular catheterization	Reward function	Motions	✓	×	×	×	✓	✓
	[150]	DDPG	Endovascular catheterization	Catheter states (positions) of current and past, motions of past	Motions	✓	×	×	×	✓	✓
	[151]	DQN	Transoral endotracheal	Simulated images	Motions	✓	×	×	×	×	✓
	[69]	DQN	Endovascular catheterization	Position data and video	Motions	✓	×	×	×	✓	✓
	[68]	GA3C	Neurosurgery	MRI images	Trajectories	×	✓	×	✓	×	✓
	[152]	DQN	Endovascular catheterization	Motion control commands from demonstrations	Motions	✓	×	×	×	✓	×
	[153]	MDP	Flexible needle insertion	CT images	Trajectories	×	✓	×	×	✓	✓
	[54]	UDQL	Flexible needle insertion	CT images	Trajectories	×	✓	×	×	✓	✓
	[154]	DQN	Flexible needle insertion	CT images	Trajectories	×	✓	×	×	×	✓
	[155]	A3C	Endovascular catheterization	Aortic arch model	Trajectories	×	✓	×	×	×	✓
Othe	[78]	K-means and SVM	Intracardiac catheterization	Images from tip camera	Type of tissue	✓	×	×	✓	✓	✓
	[87, 121]	CNN	Bronchoscopy in the lungs	Simulated images	Airway characteristics	×	×	×	×	✓	✓
	[120]	CNN	Endovascular catheterization	Motions, operating force, simulated X-ray images	(Probability of) motions	✓	×	×	×	✓	×

Table 2.7 summarizes the publications using ML techniques in the navigation task. As schematized in Table 2.7, ML applications for FSIRs's motion planning are categorized into three broad techniques: Learning from Demonstration (LfD), RL, and computer-vision-based approaches for assisting motion planning. Among different ML algorithms appearing in the navigation task, the GMM-GMR approach is dominant, as indicated in Fig. 2.4(b). It is mainly used in the LfD paradigm for motion planning because of its capability to encode the statistical characteristics from noisy demonstration data, and make predictions of robotic motions continuously over time. From Fig. 2.4(b), when it comes to RL approaches, both value-based approaches (e.g., DQN) along with policy-based approaches (e.g.,DDPG), are popular in practice. In the following, each of these methods is described referring to their three categories.

LfD-based navigation

One of the ML approaches commonly adopted to learn human-like gestures is the LfD paradigm. In LfD, expert demonstrations are used to generate a feasible task space to automate medical navigation. Statistical models such as GMM and HMM are often adopted to encode the demonstration data, which can be recorded from distal and proximal sensors embedded in the FSIRs during manual navigation. Fitted models can work as a motion planner for FSIRs navigation. The GMM-GMR approach, which is explained in Section 2.2.2, is the most commonly used approach according to Table 2.7. The expectation Maximization (EM) algorithm is often used to optimize the parameters of GMMs.

Rafii-Tari *et al.* introduced a method using learning to create motion trajectories based on various demonstrations of catheterization procedure [55]. The two-DoF linear and rotational motions produced by experts at the proximal end of a catheter are recorded during the catheterization procedure. The demonstration is then modelled with GMM. Then GMR is derived from the GMM and is applied to generate optimum motion trajectories for a robotic catheter driver. This method is able to provide assistance to inexperienced operators. The suggested cooperative scheme integrates the operator's manual handling of the guidewire with the robot's automated catheter motions, working together step by step. Additionally, the authors introduced a different collaborative approach for robotic endovascular catheterization leveraging HMM [53]. They extended their LfD approach by decomposing the procedure into a set of primitive motions and learned the model of each primitive using HMMs. A higher abstraction level HMM is utilized to grasp the overarching structure of the desired task. These refined models are used to produce a motion sequence, identify operator actions, and forecast upcoming movements. These anticipated actions during catheter navigation are showcased on a supervision interface. Through this interface, the operator observes the robot's present and subsequent movements, determining if any adjustments are needed.

Nevertheless, the anatomical structures can differ among patients in terms of size and relative location. This aspect was not addressed in the earlier studies [53, 55], but was explored in a later research [47]. This subsequent study utilized a non-rigid registration technique to identify the warping function, aligning demonstration paths to different anatomical frameworks. Robot paths are crafted based on the motion model derived from these adjusted trajectories. Both in-silico and ex-vivo tests validated the introduced path optimization and navigation strategy.

The authors of [112] exploited a model-free algorithm, i.e., policy improvement with path integrals, for motion trajectory optimization. Human demonstrations of endovascular catheterization with a 2-DOF catheter and a 1-DOF guidewire are collected and modelled by Dynamical Movement Primitives (DMPs). DMPs are then used to initialize the policy for training. Compared with manual operation, the obtained agent improved path length, speed, max acceleration and distance between catheter tip motion trajectories and vessel centerlines. This method shows better contact force results than manual and robot-assisted operations without LfD.

Building on the foundations of studies [47, 55, 112], Chi *et al.* introduced the use of Generative Adversarial Imitation Learning (GAIL) for automating robot-assisted endovascular catheterization, specifically targeting two primary aortic arch branches: the left common carotid artery (LCCA) and the brachiocephalic artery (BCA) [52]. GAIL, trained using expert demonstrations, formulates the policy for BCA cannulation. This crafted policy is then repurposed and fine-tuned for LCCA cannulation using Proximal Policy Optimization (PPO). The proposed agent's cannulation success rate is 94.4% on BCA and 88.9% on LCCA. Another work [157] also chose GAIL to train a steerable catheter in a 3D neurosurgical simulator. Its results demonstrate that the GAIL method is fast and can securely steer flexible catheters with high accuracy and robustness. In their later paper [145], similar to [52], PPO is combined with GAIL to provide intra-operative path planning for motion control. Based on this approach, they proposed a path planning framework for steerable needles used in neurosurgery. This framework can achieve an average targeting error of 1.34 mm in position and 3.16 degrees in orientation in simulation.

In the research cited as [51], an algorithm was introduced that leverages the LfD approach to learn viable paths for a flexible robot. This assimilated knowledge is captured via GMM and can be deduced as potential paths using GMR. The diversity in the acquired trajectories aids in maintaining the tip's movement along the best path while modifying the robot's form during unforeseen organ interactions. In a subsequent research [146], they expanded this method into an online learning algorithm by incorporating the Dirichlet Process (DP)-Means for real-time clustering. This refined approach creates a statistical representation of a surgeon's innate movements during a procedure and formulates a controller within the null space.

Another two works [45], [147] also took a similar LfD approach with GMM-GMR as did in [51, 55] to learn the trajectories of the tendon-driven serpentine manipulator. In [45], the GMM-GMR approach is used to learn the reaching and targeting skills from the human demonstrations and plan motion trajectories for an IK-based controller to reproduce behaviors automatically. However, Xu *et al.* proposed to learn the motor commands directly without relying on an IK model [147]. Furthermore, outcomes from GMR serve as annotated training data, enabling the use of Locally Weighted Regression model to extract models tailored for real-time control. GMM-GMR approach is also applied on a multiple-segment flexible robots, as shown in [74]. In a manner akin to [45], the study [74] posits that GMM-GMR is designed solely to generate executable trajectories. However, the control is derived from the robot's analytical kinematics model.

For the STIFF-FLOP surgical robot, a control strategy was developed by discerning context-reward relationships from demonstrations using GMM-GMR. This GMM-GMR was then employed to enhance a context-action relationship through EM [46, 123]. Alongside routine tip steering experiments, the robot underwent a test resembling a surgical setting. Here, the surgeon steered the tip, and the robot managed a midpoint on its structure. This test showcased the robot's capacity to utilize kinematic redundancy and sidestep designated areas. Subsequent research [123, 148] adopted a comparable methodology. They aimed to transfer skills observed in octopus movements to the STIFF-FLOP robot. Motion primitives, which can be variously combined for novel motions, were deduced from octopus motions via GMM and relayed to the STIFF-FLOP robot through GMR. They also designed a self-refinement algorithm to optimize GMM parameters based on iterative reward-weighted regression rather than the common EM-based approach.

RL-based navigation

An agent in RL attempts to maximize a particular reward function by taking actions and observing the consequences of those actions in the environment. Researchers adopted RL algorithms to tackle endovascular navigation in model-free approaches. Model-free approaches are commonly used RL methods that can be widely applied to different environments since they learn the optimal policy without estimating the dynamics of the environment [158]. In [149], a Q-learning-based RL strategy was introduced to navigate a robotic catheter through two distinct simulated 2D aorta meshes. The learned navigation policy on the first mesh is used as a policy initialization for the second mesh. Navigation is also performed with DRL, which benefits from the capability of ANN. An application of DQN [69] in cardiovascular intervention is already introduced in Section 2.4.3, and the navigation experiment is a showcase for their proposed control algorithm. A high success rate of 87% for the translation from a simulator to a real robotic system was reported [69]. Another work for coronary intervention is presented in [152]. In [152], a guidewire navigation framework for coronary interventions is proposed based on a DRL algorithm Rainbow [159], which improves DQN by combining multiple techniques.

The authors also made improvements by optimizing the replay memory of DQN, setting a focus window, and utilizing segment-wise training and transfer learning. The framework was experimented with a 2D coronary phantom and a 3D artery phantom with fluid flow. The final success rate of the guidewire navigation to the main target is 98% in both 2D and 3D phantoms. The authors of [155] tried to apply a DRL algorithm, namely Asynchronous Advantage Actor Critic (A3C), to learn the autonomous navigation into a virtual model of the aortic arch. Then the A3C algorithm is implemented to navigate the catheter in this virtual model. The training results show that the reward was able to achieve stability while the loss function fluctuated. Further tuning of network parameters is still required to provide a stable path.

In [99], the researchers analyzed robotic catheter navigation using two DRL methods: DQN and DDPG. They further evaluated various DDPG versions' convergence rates in a simulated setting. However, the robustness of the system was uncertain due to the simplicity of the testing environment. A subsequent study in [150] advanced this research by steering a guidewire through a rigid two-dimensional vascular model, offering a closer resemblance to human vessel structures than their prior work. In this study, DDPG, enhanced with hindsight experience replay, directed the guidewire. This was trained in a simulation where the agent, based on present and previous catheter states and earlier motions, determined the subsequent motion, comprising both translation and rotation commands.

Works emerging with RL methods in other clinical applications are studied as well. [151] developed a navigation policy based on Deep Q Reinforcement Learning Neural Networks (DQNN) that utilizes images from a monocular camera mounted on the tip of a snake robot for tracheotomy. The system is said to serve as an assistive device for medical personnel to perform endoscopic intubation with minimal human input.

In addition to its use with catheters or concentric tubes operating within lumens, RL has been employed for the insertion of flexible steerable needles. As demonstrated in [54], a supervised approach is taken for both path and motion planning of steerable needles using Universal Distributional Q-learning (UDQL). This agent, trained in simulated 2D and 3D settings, can execute insertions targeting multiple points from a singular entry point, augmented by hindsight experience replay. This method produces a distribution indicating potential risks, allowing professionals to manually adjust and refine the model based on their assessment, facilitating transfer learning. The proposed approach performed better than the deep double Q-learning network, which is a variant of DQN, in terms of steering accuracy and avoidance probability. In their previous work [153], which is not based on RL but on MDP, they proposed a robust path planning algorithm to provide secure and optimal motion planning. The navigation challenge is framed using an MDP method, where uncertainties are treated as variables with undetermined distributions. The proposed method outperforms the traditional MDPs in success rate and avoidance probability. In the study by [154], DQN was employed to master the insertion of a flexible needle with 2 DOFs. The reward system was based on the proximity between the needle tip and the target, while penalties were imposed for contact with surrounding tissues. The GPU-based A3C algorithm (GA3C) is also applied to plan paths for a steerable catheter in [68]. It delivers the best paths considering obstacle avoidance and kinematic limits. The outcomes demonstrated leading-edge results in terms of navigating around obstacles, trajectory smoothness, and computational speed.

Other methods

In various medical settings, flexible robots are often equipped with a camera at their tip, providing live visual data of the immediate surroundings. This visual feedback can be used to guide the robot's navigation. These methods are classified as computer vision-based navigation.

Fagogenis *et al.* introduced an algorithm for navigation relying solely on visual cues [78]. Using a limited image dataset, they fully trained K-means and SVM image classifiers. These classifiers are adept at differentiating between ventricular walls or blood and the bioprosthetic aortic valve. Based on the tissue the FSIRs interacts with, the system can toggle between two navigation strategies to perform intricate clinical procedures.

In addition to SVM classifiers, ANNs, particularly CNNs, have gained popularity for visual-assisted FSIRs navigation. In studies by [87, 121], CNN designs were formulated to instantly determine the location of the bronchoscope using information from a camera at its distal end. The subsequent closed-loop control was visually guided. These CNNs are trained to recognize features of the observed airways. This data is then relayed to the robotic bronchoscope's motion controller, which then selects the optimal trajectory from a set of predetermined airways. With the proposed navigation method, the robot successfully reached the target in the lung in 19 out of 20 trials. A DL methodology that uses CNNs for steering an endovascular robot draws insights from X-ray images of the present clinical condition. In the study by [120], Zhao *et al.* introduced a framework that merges simulated X-ray visuals from an exterior camera and internal operational forces to guide navigation decisions. The training data, which comprises operational actions, operational force, and medical state images, is used to train two CNNs, *i.e.,* a 2D CNN that learns from medical images to estimate the probability of action and a 1D CNN that utilizes operating forces to identify the force pattern. This navigation framework achieved a success rate above 84% on a mixture of data from two cases.

2.5 Discussions on the survey of the ML used in FSIR

The ML approaches have been successfully applied in FSIRs tasks involving perception [44, 124, 127], modeling [63, 64, 76, 91], control [75, 105, 115, 120], and semi-autonomous/autonomous navigation [47, 52, 55, 78]. Figure 2.5(a) reflects the frequency of FSIRs-related papers using ML techniques by year. The number of papers is steadily increasing with a peak in 2019. Figure 2.5(b) summarizes the exploitations for the four categories of ML approaches. The number of traditional ML algorithms used can be seen to fluctuate from year to year and its appearance remain stable from 2013 to 2020. This may be related to the fact that traditional ML approaches remain the most suitable choices when working with small datasets, very common in the FSIR context, where DL techniques cannot be trained effectively. The application of DL has experienced rapid growth since 2017. This is consistent with other scientific fields and explained by a true DL boom that started with Alpha Go in 2016. In addition, the increased accessibility of open source DL libraries (e.g., Tensorflow, first released in November 2015, and Pytorch, first released in September 2016) has fed this booming trend. As is demonstrated in Fig. 2.5(b), the application of RL have also increased from 2017 onward. Since 2017, 20 out of 27 RL-relevant articles have utilized DL techniques. The DRL has benefited from the growth in ANN. Despite the above-mentioned growth in popularity, applying different ML methods to FSIRs still suffers from some common limitations/challenges. We want to discuss some common hindrances, which are found from the reviewed literature, as well as some open issues that are not often mentioned in recent researches.

Performance gaps among simulation, bench-top experiment and in-vivo test

A well-accepted procedure to validate algorithms on autonomous robots is to firstly test the robotic functionality in computer-based simulators and bench-top synthetic phantoms before carrying out in-vivo animal/human trials [69, 86, 108]. Prior to a clinical application, safety, ethical and legal guidelines need to be considered.

However, the discrepancies between a simulated and a bench-top environment, called the “reality gap”, can lead to significant shortcomings when a ML model trained on simulated data is deployed in a bench-top environment. These differences are due to modeling errors since it is highly challenging to model contact forces, friction, tool-tissue interaction, sensor noise, and lighting conditions, all leading to performance decay of ML methods when the training and test environment vary. Therefore creating surgical simulations is a widely studied topic in the literature which requires realistic and real-time modeling of soft tissue response to tool-tissue interactions, in addition to realistic rendering [160].

Various studies have attempted to solve the reality gap by introducing perturbances in the environment or focus on domain randomization [161]. Domain randomization techniques try to generate large volumes of simulation data by considering the sim-to-real differences in a virtual environment [162, 163]. In the reviewed works of FSIRs, randomized noises are added to the simulation environment, so as to make the simulated experiment more complicated and realistic [51, 69, 98, 112]. Another solution is generating synthetic data that are close to real data. Some recent works have used Generative Adversarial Networks (GANs) to construct a mapping between simulated and real domains, then to create synthetic data, such as [133, 164, 165].

One can expect that another gap exists between bench-top and in-vivo experiments. Here bench-top experiments refer to either synthetic phantom or ex-vivo tissues. This gap has been less reported because there have been few works applying ML techniques in in-vivo experiments [78, 142]. In-vivo environments are considered to be more complicated because the robot has to interact with different body fluids and soft tissues. Another challenge comes from physiological movements, such as breathing and heartbeat, which make in-vivo environments dynamic. Thus, maintaining precise control of robots is more challenging in in-vivo environments.

Limited interpretability of ML algorithms

In Section 2.2, the interpretability of ML and some interpretable ML models are described. However, the limited interpretability of the ML algorithms are not yet discussed. Interpretability is defined as *"the degree to which a human can understand the cause of a decision"* [166]. A ML model with concise mathematical or statistical expression, such as a linear regression model, makes it easier to comprehend why such a model makes certain predictions or decisions. The limited interpretability of some ML models raises safety concerns on the application of these ML techniques on FSIRs. One typical example is ANN, which is often viewed as a "black-box" because it is difficult for a human to follow the data stream from the raw input to the network output. In addition, the knowledge that ANN learns is stored in hidden layers. The parameters of an ANN such as weights and biases that a human can directly observe do not contain any physical meaning. This information is thus too complex for a human to follow. Therefore, the use of ANN with FSIRs may raise doubts on reliability and acceptability issues to clinicians.

To increase the understanding of ANNs, ablation study has been used when applying ML algorithms on the FSIRs [85, 138]. Through an ablation study, one can determine the contributions of each part in an ML algorithm to the overall performance, or justify the selection of the ANN hyperparameters.

The limited interpretability of ML algorithms also increases the uncertainties of an FSIRs system. In safety-critical areas such as MIPs, it is crucial to adhere to high safety standards as errors in robotic control can lead to hazardous situations

[167]. However, it is difficult to guarantee safety in all circumstances when dealing with complex computer-controlled systems such as FSIRs [168], especially when ML techniques are incorporated (even experienced surgeons or interventionists are not flawless).

Due to the uncertainties involved in ML models, providing guarantees of safe behavior in a trained model is highly challenging [169]. To mitigate the risks caused by uncertainties, recent regulatory measures in the high-risk category, such as surgical robots, are directed towards involving human supervision [167]. Hence, systems integrating ML techniques may consider including humans in the loop in earlier designing phases [167], which could further enable shared control. The human-involved shared control can be done by, e.g., visualization in real-time of the system's status or predictive display of future actions the system intends to pursue before actually executing those actions [170]. It is also possible to evaluate the risk of future actions by algorithms and manage the risk with human supervision, as implemented in [54]. Another example of shared control is that manual and autonomous control are switched in different phases of the MIPs, depending on the complexity of the tasks[55]. Moreover, some studies have proposed safe-reinforcement learning using formal verification methods for robot-assisted MIPs [171]. Other techniques for safe-reinforcement learning that have not yet been applied in the surgical domain are reviewed in [172].

Data issues

In terms of data collection, unlike image recognition or natural language processing, where huge labeled data sets are open to the public, medical data for FSIRs research, such as endoscopic images, is typically limited in size. It is often difficult for researchers to collect more clinical data from external data centers because sharing medical data still has privacy concerns. As an attractive solution to overcome the shortage of medical data, federated learning has been embraced by the medical image community. Federated learning utilizes data from individual data sites to train a global model without sharing the data directly, so as to improve the robustness and performance of the global model while protecting data privacy.

Another problem is the bias in data because surgical specialties and employed instruments may be quite dependent on the center where data was gathered. Data lacking the variability across centers may result in potentially ill-suited systems and misdiagnoses [173]. For instance, in the clinical studies of [173], excluding data of African-Americans led to the misclassification of some patients as pathogenic. In practice, to enrich training data for ML, artificially manufacturing synthetic data has been adopted in clinical studies [174]. In the context of FSIRs, synthetic data is often used to enlarge datasets with real sensor data [88, 133, 165]. However, creating assessment criteria for objectively evaluating synthetic data is still an open question [175].

Different vision between engineers and clinicians

Yet another problem is the difference between metrics of algorithms and clinical needs. Researchers may be excited to see ML model performance metrics improve, but this does not necessarily result in an improved clinical outcome, which matters for clinicians (and patients). From the clinicians' perspective, performance of ML algorithms may not be the most convincing factor to use ML. Sometimes the improvements in performance metrics are at the expense of changing the clinicians' customs. Clinicians may also have concerns about the applicability of ML algorithms in complex clinical applications. ML algorithms' ability to justify its outputs and help clinicians understand the output has been generally believed to be crucial to establish clinicians' trust in ML [176]. To eliminate barriers in vision, researchers have to collaborate more closely with clinicians and evaluate the results from the clinician's perspective from bench-top to in-vivo experiments.

Ethical and liability aspects

Employing ML algorithms in conjunction with FSIRs poses challenges not just for researchers but also necessitates a wider debate concerning ethics and liability. From the ethical side, informed consent is a principle in healthcare, but it could be a challenge when ML is involved in clinical practice. As we discussed before, due to the limited interpretability of some ML algorithms, clinicians may feel it hard to understand the reason for ML's outputs, let alone inform and educate the patients about the complexity of ML used by the devices. In addition, liability for ML-involved clinical devices such as FSIRs is another challenge. Who will be liable if a FSIRs with deployed ML algorithms makes a mistake? The broader ethical and legal issues related to AI in healthcare have been elaborated upon by [84].

2.6 Conclusions on the survey of the ML used in FSIR

This chapter reviews and discusses current applications and research activities of ML algorithms in the context of FSIRs. ML has played an increasingly important role in different tasks of FSIRs, such as perception, modeling, control and navigation. From the perspective of interventional tasks, this chapter aims to sketch the big picture showing how different ML algorithms are gaining popularity in various scenarios, thereby providing a good indication of how advances in ML could boost the use of FSIRs in clinical procedures. From the perspective of ML algorithms, readers can gain insights into which functionalities ML could perform. This provides guidance for researchers when weighing the use of ML versus traditional model-based techniques for FSIRs in MIPs. This chapter also described an analysis of a possible workflow or interaction between ML algorithm, clinicians and FSIRs. ML algorithms may enhance the autonomy level of FSIRs by engaging in various phases of the MIPs. The clinicians could focus on high-level tasks when working with ML-based FSIRs. Meanwhile, the expert demonstration may help the ML algorithm learn skills from clinicians, so as to improve the autonomy level of FSIRs. In this manner, FSIRs hold the potential to be more intelligent with the use of ML over time.

Advances of autonomy in FSIRs still face challenges due to the limitations of ML and clinical factors. Much work is needed before these techniques can be used in real clinical practice. As a matter of fact, the validation of the vast majority of the works presented in this chapter only took place in in-vitro experiments. Data of high quality is always important to train ML models, yet it is not easily accessed and tricky to process. Thus, the challenge of obtaining data amid growing concerns about privacy is an issue that urgently needs addressing, for example, with federated learning. Also, some popular ML algorithms often behave as a black box. It remains an urgent objective in increasing the interpretability of ML. Moreover, this is likely to be crucial if one wants to convince the developers but, more importantly, the clinicians to put their trust in said ML algorithms. Improving interpretability is crucial to building trust in ML, thereby enabling more effective use of these algorithms in FSIRs systems in clinical practice.

This chapter addresses sub-objective 1 by surveying the latest applications of ML in FSIR, thereby laying the groundwork for the use of ML techniques in addressing control and sensing challenges in robotic catheters. This approach is in alignment with the overarching goal of the thesis. Moreover, the survey provides valuable insights into algorithmic choices, which benefits the application of ML techniques in the following chapters. In Chapter 4, a MLP network is used to map the wavelength shifts of FBG sensors to the curvature and the angle of the bending plane of the continuum robot. MLP is chosen due to its capability to model nonlinear systems and its ease of implementation. Furthermore, in Chapter 5 and Chapter 6, a type of Recurrent Neural Network namely Long Short-Term Memory network was chosen to address hysteresis in robotic catheters. Understanding the characteristics of various ML algorithms and their most practical applications enables us to identify the most suitable algorithm for a specific case, thereby reducing the time and effort spent in experimenting with different algorithms.

Chapter 3

Comparative analysis of interactive modalities for intuitive endovascular interventions

This chapter is based on the submitted journal article:

D. Wu*, Z. Li*, M. H. D. Ansari, X. T. Ha, M. Ourak, J. Dankelman, A. Menciassi, E. De Momi, E. Vander Poorten,“Comparative Analysis of Interactive Modalities for Intuitive Endovascular Interventions”.

Endovascular intervention is a minimally invasive technique for treating cardiovascular diseases. Fluoroscopy, known for its ability to visualize catheters in real-time, is commonly employed to ensure accurate catheter placement. However, fluoroscopy exposes patients and physicians to ionizing radiation and offers no depth perception due to its 2D nature. This chapter aims to overcome some of these limitations by employing teleoperation and 3D visualization. This chapter presents three interactive control modalities to steer robotic catheters and investigates their effects in human-in-the-loop robot-assisted endovascular intervention. The developed methods employ a gamepad, a standard 2D monitor, and an Augmented Reality (AR) Head-Mounted Display (HMD). In this work, a Microsoft HoloLens 2® is utilized as an advanced HMD AR headset featuring 3D visualization. An *in-vitro* user study was conducted to compare three interactive modalities: 1) Mode GM - a gamepad with a standard 2D monitor, 2) Mode GH - a gamepad with a HoloLens, offering a 3D visualization, and 3) Mode HH - a HoloLens serving both as input device and as visualization device. A robotic endovascular catheterization system was set up for experimental validation. User performance was subjectively and objectively assessed. The use of a gamepad for steering and a HoloLens for visual feedback (Mode GH) scored the best on all subjective metrics except mental demand. Mode GH outperformed the other modalities also in regard to objective metrics showing median tracking errors of 4.92 mm, median targeting errors of 0.93 mm, and median durations of 82.3 s in the conducted *in-vitro* study. Mode GH offered improvements of 8.7%, 15.5%, and 10.9% respectively over Mode GM, and 4.8%, 37.2%, and 31.9%, respectively with respect to Mode HH. The study conclusively demonstrates the superiority of using a gamepad to teleoperate catheters. Such a teleoperation approach offers several advantages, including its user-friendly nature, superior responsiveness (when compared to HoloLens's hand gesture recognition), and portability. Furthermore, the study showcases the practicality of using HoloLens to receive 3D visual feedback, which enhanced the visualization in catheterization.

This chapter directly aligns with Sub-objective 2 of the thesis, focusing on exploring intuitive interaction modes for teleoperated endovascular robotic systems. Developing an intuitive mode for steering the catheter, and combined with reliable visual feedback, can significantly enhance the safety of interventions.

3.1 Introduction

Throughout the catheterization procedure, physicians can employ real-time imaging techniques, such as fluoroscopy, to visualize the heart and blood vessels [177, 178]. This facilitates the identification of lesions, navigation itself, and immediate evaluation of treatment effectiveness. However, fluoroscopy exposes both patients and physicians to ionizing radiation. In particular, physicians who are repeatedly exposed to radiation face a higher risk of cancer and cataracts [179]. Furthermore, the two-dimensional (2D) nature of fluoroscopic images prevents depth perception [180, 181], complicating the precise maneuvering of catheters within intricate vessels.

Compared to current practices, teleoperation has emerged as an advantageous approach for catheterization procedures. Its primary benefit is to reduce radiation exposure for physicians. Teleoperation also increases precision, reduces human error, and facilitates remote surgeries. However, it is important to note that this control method lacks autonomy and relies entirely on the input of the operators. In addition, advances in visualization technology have introduced sophisticated techniques such as Augmented Reality (AR), Virtual Reality (VR), and Holographic Imaging, all of which provide enhanced visual assistance. Considering the multitude of methods available for both teleoperation and visualization, this chapter conducts an investigation into the effectiveness of these various interactive modes. This investigation is guided by a comprehensive set of experiments and evaluation metrics.

One approach to tackle the challenges mentioned above is 3D imaging and visualization. A variety of methods have been used to produce a 3D model of the vessel. Some groups employ 3D rotational angiography [182]. Others propose to fuse the pre-acquired 3D images (e.g., from CT or MRI) with real-time fluoroscopy [180, 183]. These methods can provide improved depth information, but radiation exposure remains a concern. Non-radiative imaging modalities, e.g., OCT [184] and IVUS [185], could also be employed to produce some sort of depth information during catheter-based procedures. By utilizing the aforementioned techniques, 3D reconstruction of 2D medical images could be made available. However, the optimal use of this 3D content still requires further investigation. Displaying these 3D volumetric data on a 2D screen is a cheap and common method. However, 2D screens do not manage to convey the third dimension very well [186]. Moreover, 2D screens hinder direct interaction with the image [187]. AR systems present an appealing alternative interface by overlaying 3D rendered images onto the physician's field of view during the procedure. This technology could improve depth perception and spatial understanding, facilitating more accurate instrument placement and navigation [188]. It should be noted that while many 3D reconstruction methods have been introduced previously, the primary objective of this study is to investigate the optimal presentation of the 3D content, with different display techniques, rather than investigating intraoperative 3D reconstruction methods themselves.

Along these lines, an image guidance system using AR visualization is proposed for transcatheter procedures in [189]. The 3D hologram displayed in AR was the preoperative 3D model reconstructed from CT scans. The catheter tip position is derived from 2D segmentation of intraoperative fluoroscopy images and registered to the 3D model. Although this method improves visualization, radiation exposure from fluoroscopy remains a concern. Palumbo *et al.* explored the use of AR for radiation-free catheter navigation [190]. In their work, the guidewire tip pose is determined by EM trackers. However, their work only introduced a registration method that aligns the EM tracker's position with a holographic marker. Guidewire or vessel phantoms were not used to investigate in detail the influence of AR on vessel navigation. A related approach is presented in [191] for EVAR, where the catheter tip position is also obtained using an EM tracker. This work merely represents the tip position as a sphere. This representation neglects the tip orientation, which is crucial for assessing whether the catheter is in a hazardous pose. Furthermore, the paper only describes the concept and implementation without providing quantitative results on the effectiveness of the assistance offered by visualization VR.

Linte *et al.* achieved catheter tip visualization in AR through ultrasound imaging [192]. The work evaluated the effectiveness of AR visualization by having three novices use a commercial cable-driven manual catheter to target four specific locations within a heart phantom. The results indicate that AR visualization could enhance target accuracy compared to 2D ultrasound images, while the assistance provided by AR in vascular navigation has not yet been validated. Apart from the limitations mentioned above, physicians will need training and time to become comfortable with AR systems, and extra research is needed to achieve optimal integration of AR into the clinical workflow. Upon a review of the previous literature, it was found that no studies investigated *in-vitro* the potential of AR for endovascular intervention with robotic catheters. A possible explanation for the lack of such studies is the limited availability of robotic catheter platforms, the recent availability of high-quality AR headsets, and the complexity to integrate all these components.

Teleoperated catheter navigation can also significantly reduce radiation exposure [193]. This technology could also potentially allow physicians to perform intricate procedures with enhanced precision and reduced physical strain. Teleoperation can be accomplished using a variety of control input devices. Researchers have explored the use of joystick-like controllers [194, 195], an in-house developed haptic joystick [30], and custom-made input devices [196] for teleoperating catheters.

Table 3.1: Comparison with previous work

References	**Procedure**	**Device**	**Manual/Motorized catheter delivering**	**AR device**	**Intraoperative feedback on catheter pose**	**Evaluation methods**	**Validation through vessel navigation**
[189] Liu *et al.*	transseptal puncture	Destino$^{\text{TM}}$ Reach catheter	manual	HoloLens	dual fluoroscopic images	image registration accuracy	yes, 1 user
[190] Palumbo *et al.*	structural interventional cardiology	guidewire	manual	HoloLens 2	EM tracking system	registration accuracy	no
[191] Garcia *et al.*	endovascular aneurysm repair	stent graft catheter	manual	HoloLens	EM tracking system	system feasibility	yes but no quantitative evaluation
[192] Linte *et al.*	atrial ablation	Freezor$^{\text{TM}}$ catheter	manual	2D monitor	ultrasound images and EM tracking system	targeting accuracy	yes, 3 users
This work	endovascular intervention	custom-made robotic catheter	motorized, teleoperation with three modes	HoloLens 2	EM tracking system	tracking, targeting accuracy + NASA-TLX	yes, 9 users

Additionally, the market offers a diverse range of control input devices for human-machine interfaces. These include haptic devices such as the Touch and Touch X haptic devices by 3D Systems, which can provide force feedback. Foot pedals offer hands-free operation [197], while traditional joysticks [198], commonly used in aviation, offer straightforward control. Trackballs are another option, providing precise cursor control through a stationary, rotatable ball [199]. The user-friendly nature and widespread familiarity with joystick-like controllers can facilitate a smoother transition from gaming to interventions and a more rapid learning curve for physicians using teleoperated catheter systems. This can subsequently result in enhanced procedural efficiency, patient safety, and outcomes. The custom-made input device designed in [196] consists of a tube and several sensors, resembling a catheter. This enables physicians to manipulate the input device more intuitively using familiar actions such as rotation and pushing, similar to how they would handle a catheter in current practice.

In this study, two visualization methods for 3D content are investigated: 1) a standard 2D monitor and 2) an advanced 3D visualization technique using AR. One aim of this work is to determine which visualization approach is more effective at this stage in a first *in-vitro* study. The other aim of this work is to primarily compare the use of a gamepad and a Head-Mounted Display (HMD) as input devices. To the best of the authors' knowledge, this is the first time that an HMD has been utilized as a control input device for catheter steering.

This chapter presents a user study investigating the effects of different interactive modalities in human-in-the-loop robot-assisted endovascular intervention with visual feedback. The main contributions of this chapter are:

- design of an AR interface in HMD for enhanced visual feedback and catheter steering, implementing gamepad-based teleoperation to steer a robotic catheter system, integrating stand-alone components e.g. robotic catheter, catheter driver, HMD and gamepad into a robotically-assisted endovascular interventional system;
- *in-vitro* user study with various combinations of control and visualization devices, involving nine participants with diverse levels of experience in HMD, gaming, and steerable catheters;
- a detailed performance analysis and discussion comparing the different interactive modalities.

This chapter is organized as follows: Section 3.2 provides an overview of the hardware components used in the system and their integration. Section 3.3 explains the experimental design and the performance metrics used to evaluate the efficacy of catheter navigation. Section 3.4 presents the results, accompanied by a discussion, while Section 3.5 offers conclusions and describes future work.

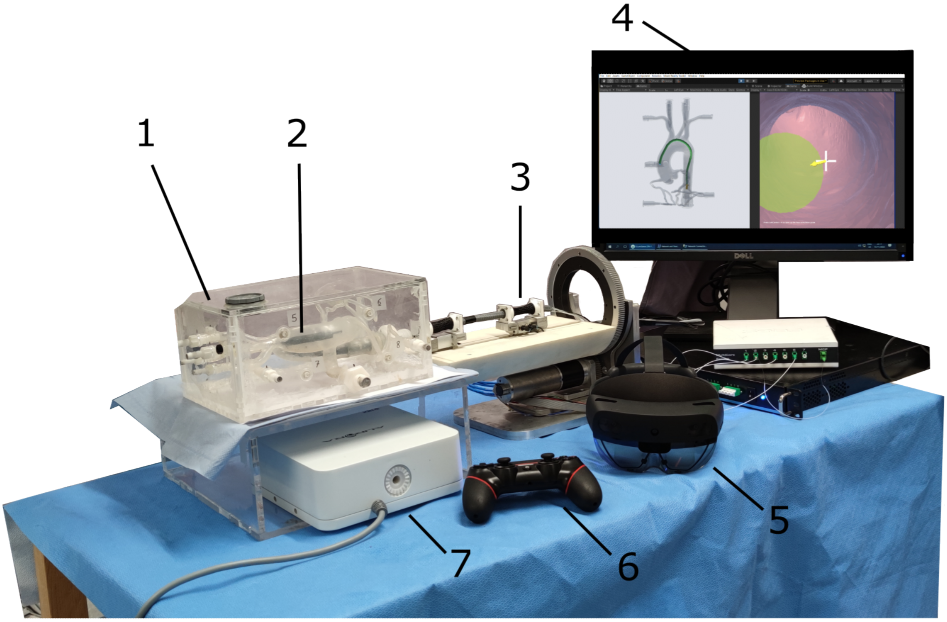


Figure 3.1: Experimental setup for *in-vitro* user study: 1. silicone aortic phantom; 2. pneumatically-driven robotic catheter; 3. catheter driver; 4. standard 2D monitor; 5. Head-Mounted Display (HMD); 6. wireless gamepad; 7. electromagnetic field generator.

3.2 Experimental setup and methodology

3.2.1 System components

To investigate various interactive modalities for catheter navigation, an experimental setup for endovascular intervention was developed (as shown in Fig. 3.1). It consists of the following components:

silicone aortic phantom an aortic phantom (T-S-N-002, Elastrat Sarl, Geneva, Switzerland) made of silicone, replicating key features such as the descending aorta, the aortic arch, the aortic root, and several coronary arteries.

robotic catheter an in-house developed robotic catheter with a 50 mm bendable Nitinol segment, equipped with a pattern of equally spaced slots produced by EDM [200]. These slots make the segment bendable. The remaining section, approximately 1 m in length, is composed of a predominantly passive flexible

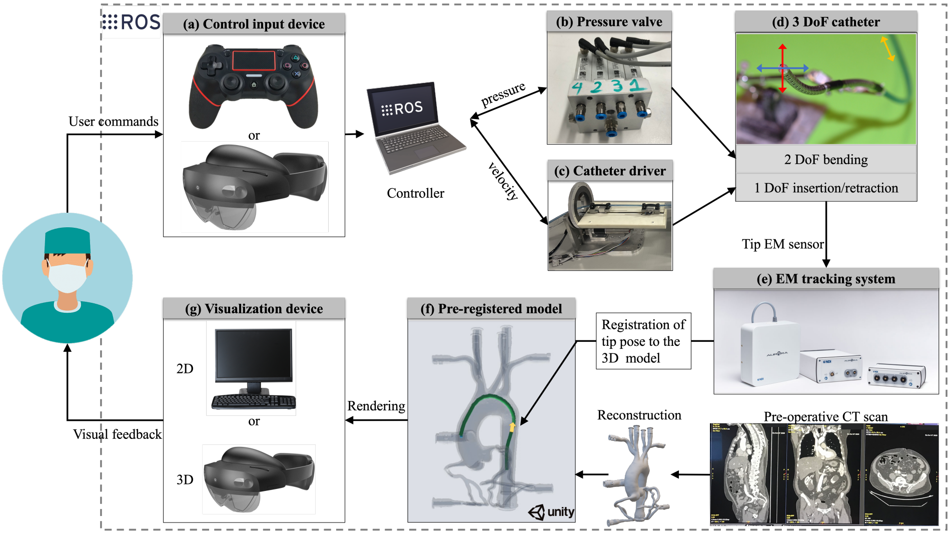


Figure 3.2: Schematic of the advanced human-in-the-loop vessel navigation system with multiple interactive modalities, including: (a) a gamepad or an HMD as control devices; (b) a pneumatic valve with four output ports receives control commands from ROS; (c) the catheter driver, operated through velocity control, regulates the insertion and retraction of the catheter; (d) the PAM-driven catheter, having 2-DoF and a 60 mm active bendable segment; (e) an EM tracking system localizing the catheter tip, whose pose is then registered to the mesh frame of the 3D reconstructed model; (f) the virtual 3D aortic model reconstructed utilizing high-resolution CT images, along with the guidance path and the catheter tip pose registered and rendered in this virtual model frame; (g) visual feedback that users receive either through a standard monitor (2D) or via a HoloLens (3D).

plastic tube. With a diameter of 7 mm, the catheter is well suited for navigation within the aorta, as the typical diameter of the aorta ranges between 20 and 35 mm [201]. The bendable Nitinol segment of the robotic catheter features four PAM, organized into two antagonistic pairs [200]. Each PAM has one end connected off-center to the catheter tip. When pressure is applied to the muscle, its length decreases, generating a pulling force on the catheter tip, forcing it to bend. By simultaneously controlling two antagonistic pairs of PAMs, the catheter achieves a spatial 2-DoF bending motion. A 6-DoF EM sensor (Aurora®, Northern Digital Inc., Canada) is glued to the center of the catheter tip, enabling precise tip localization.

catheter driver the catheter driver [202] is designed for axial rotation about the longitudinal axis of the catheter and for 1-DoF catheter insertion or retraction. The device is based on two sleeve-based grippers. One end of the sleeve remains stationary, while the other end is attached to a pneumatic piston. As pressure increases, the sleeve expands in length, subsequently reducing its diameter and ensuring a firm grip on the catheter body. The two grippers operate alternately, with one gripping as the other releases. This synchronized operation allows continuous catheter translation over a long stroke.

Head-Mounted Display (HMD) Microsoft HoloLens 2®, simply referred to as "HoloLens" in the following, is an advanced HMD AR headset featuring 3D visualization. Its advanced display technology projects holograms, which are 3D virtual objects, into the user's field of view. These holograms blend seamlessly with the real world. The HoloLens also has sophisticated hand-tracking capabilities for both hands. The HoloLens hand-tracking system enables smooth interactions for users, allowing them to select and position holograms using direct touch, as if interacting with tangible objects in the real-world space. Furthermore, virtual rays projected from the center of the user's palm function as extensions of their hands. This design allows users to interact effortlessly with holograms located beyond their immediate physical reach. The wireless design of the HoloLens promotes unlimited movement without the burden of external cables. Voice control is another feature of the HoloLens. Despite its advanced features, the headset weighs only 3.28 kg.

gamepad the controller (Yues, Dublin, Ireland) is a wireless input device for gaming. The controller is equipped with four buttons on the left, two central thumbsticks, and four additional buttons on the right.

electromagnetic field generator an EM field generator (Northern Digital Inc., Canada) is placed beneath the phantom. When the EM sensor enters the electromagnetic field produced by the field generator, it induces a small current within the sensor. This current is then converted to the corresponding positions and orientations of the EM sensor. However, the presence of electromagnetic materials within this field tends to distort the accuracy of these measurements. For this reason, the catheter driver's metal components are located outside the generated magnetic field so as not to disturb it. The 6-DoF EM sensor (Aurora®, Northern Digital Inc., Canada) embedded in the catheter tip allows tip pose measurement at 40 Hz.

software architecture communication between various devices is facilitated through ROS 1 [203], with each device functioning as a distinct node within the system.

Figure 3.2 presents the control scheme for human-in-the-loop catheter navigation with visual feedback. Users can interact with either a gamepad or a HoloLens to

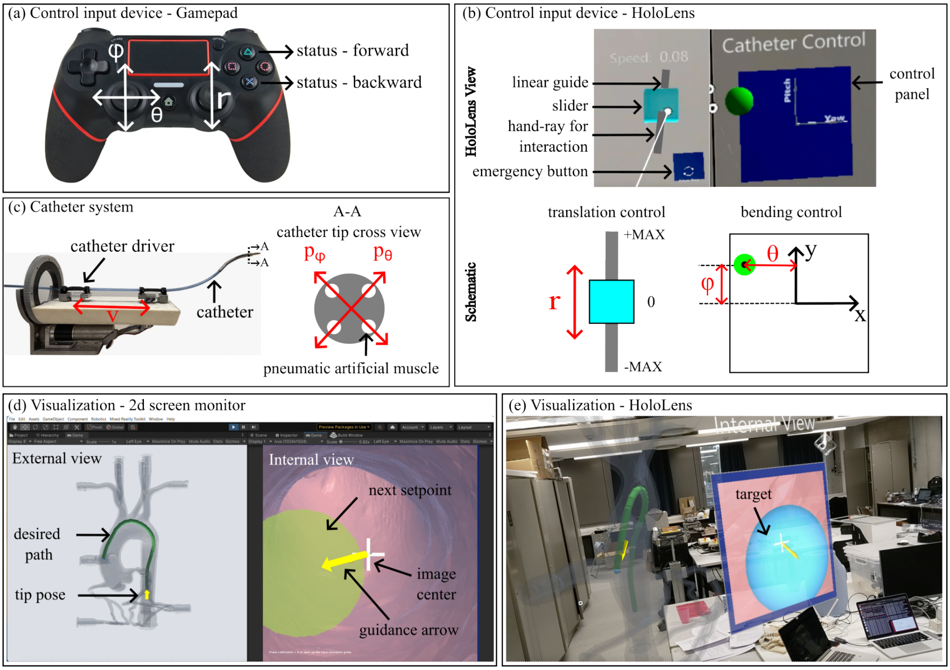


Figure 3.3: User interfaces for input devices and visual feedback: (a) a gamepad as a control input device; (b) a HoloLens serves as a control input device, featuring hand gesture recognition capabilities; (c) a schematic illustration outlining the three mapping relationships corresponding to the 3-DoF for the catheter system; (d) 2D visualization using standard monitor; (e) 3D visualization using HoloLens: Users can view the holograms from various angles by physically moving around in real-world space.

generate control commands, while obtaining 3D visual feedback from an HoloLens or 2D visual feedback from a standard 2D screen.

3.2.2 System integration and user interface design

This subsection explains how the different parts are integrated into a endovascular robotic system and describes the design of the user interface.

The functionalities of the various thumbstick or buttons are illustrated in Fig. 3.3(a). By using a ROS joystick driver library, the bending angle of the thumbstick can be read as a floating value between 0 and 1, while the button toggles between two distinct values: 0 and 1. Users can control the 2-DoF bending by directly mapping the bending angle of the left thumbstick (ϕ, θ) to the pressure applied to PAM (p_ϕ, p_θ), so that p_ϕ is proportional to ϕ, and p_θ is proportional to θ.

The maximum pressure applied to PAMs is set at 6 bar. The right thumbstick is employed to regulate the catheter driver's insertion/retraction motion speed. The bending angle of the thumbstick r is proportional to the translation speed of the catheter driver v, where the catheter driver is operated using velocity control and the velocity of the catheter driver v is proportional to the bending angle of the thumbstick r. The maximum translation speed was set to 5 mm/s. The PS (Δ) and OPTIONS (X) buttons on the right (see Fig. 3.3(a)) are used to switch between the insertion and retraction modes. The choice to select between forward and backward motion commands rather than mapping the gamepad's bidirectional motion to forward and backward motion was found to help prevent the catheter driver from switching continuously between forward and backward motion, making this a more deliberate/controlled choice by the user. The thumbstick is equipped with a spring-like automatic return mechanism, which may lead to input fluctuations around 0 when released abruptly. As a result, the catheter driver may rapidly alternate between forward and backward movements in a short period of time. The method of consciously selecting between movement directions is adopted as an alternative to the establishment of a deadzone where the catheter driver's speed defaults to zero. It may be worth comparing both methods in future user trials.

Figure 3.3(b) illustrates the user interface design of the HoloLens when used as a control input device. Users can engage with virtual buttons or sliders shown by the HoloLens in a manner akin to interacting with actual objects, to send control commands for steering the catheter. The speed, v, of the catheter driver is regulated by the translational distance of the slider, r. An emergency button has been incorporated to allow users to quickly stop the catheter's translational movement in case of an emergency. The blue control panel is displayed on the right side in the upper image of Figure 3.3(b). The green sphere can be moved in this 2D plane. The coordinates of the sphere are measured and proportionally converted into pressures. These pressures are then transmitted to the Pneumatic Artificial Muscles (PAMs) in two orthogonal directions, denoted as p_ϕ and p_θ. Users can reposition various holographic components in the HoloLens view by grasping and dragging them, allowing them to customize the layout to their optimal comfort level. Moreover, users can engage a hand-ray, a common remote control feature in HoloLens, that extends from their palm towards the holographic object. This allows them to interact with the holographic object even when the object is situated beyond their physical reach.

The control commands generated by the gamepad or HoloLens are sent to the pressure valves, which adjust the bending of the catheter, and to the catheter driver, which manages the translational movement of the catheter, as shown in Fig. 3.3(c). This process allows for precise control over the robotic catheter's 3-DoF.

An EM tracker, attached to the tip of the catheter, captures its pose within the magnetic field created by an EM field generator. The tip pose is expressed through the transformation matrix ${}^{\boldsymbol{E}}\boldsymbol{T}_{\boldsymbol{C}}$ that registers the tip frame $\{\boldsymbol{C}\}$ to the EM frame $\{\boldsymbol{E}\}$, with ${}^{\boldsymbol{E}}\boldsymbol{T}_{\boldsymbol{C}}$ determined by the EM measurement. A 3D aortic mesh model is reconstructed from CT images. To correctly represent the catheter pose in the aortic model $\{\boldsymbol{M}\}$, namely ${}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{C}}$, registration ${}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}$ must be performed.

$$ {}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{C}} = {}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}\,{}^{\boldsymbol{E}}\boldsymbol{T}_{\boldsymbol{C}} = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0} & 1 \end{bmatrix} {}^{\boldsymbol{E}}\boldsymbol{T}_{\boldsymbol{C}} \tag{3.1} $$

Ten markers are employed to calculate the transformation matrix ${}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}$ that registers the EM frame $\{\boldsymbol{E}\}$ to the virtual mesh model frame $\{\boldsymbol{M}\}$. These markers are strategically placed in the outer casing of the silicone aortic phantom (see Fig. 3.4). Two datasets are involved in this process: one comprises the positions of the ten markers ${}^{\boldsymbol{M}}\boldsymbol{p} = \{{}^{\boldsymbol{M}}\boldsymbol{p}_i, i = 1, 2, ..., m\}$ within the virtual model frame $\{\boldsymbol{M}\}$, and the other contains the positions ${}^{\boldsymbol{E}}\boldsymbol{p} = \{{}^{\boldsymbol{E}}\boldsymbol{p}_i, i = 1, 2, ..., m\}$ of the ten markers represented in the EM frame $\{\boldsymbol{E}\}$, where $m = 10$.

The optimal transformation matrix ${}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}$ is determined by a SVD method, following [204]: (i) identify the centroids of both datasets, ${}^{\boldsymbol{M}}\boldsymbol{p}$ and ${}^{\boldsymbol{E}}\boldsymbol{p}$; (ii) shift both datasets to the origin, then compute the optimal rotation $\boldsymbol{R} \in \mathbb{R}^{3\times3}$ using the SVD method; (iii) calculate the translation vector $\boldsymbol{t} \in \mathbb{R}^3$. The objective function is defined as follows:

$$ \min_{{}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}} f = \sum_{i=1}^{m} ||{}^{\boldsymbol{M}}\boldsymbol{p}_i - {}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}\,{}^{\boldsymbol{E}}\boldsymbol{p}_i||^2 \tag{3.2} $$

RMSE, as defined by (3.3), was calculated based on the ten registration markers.

$$ RMSE = \sqrt{\frac{\sum_{i=1}^{m} ||{}^{\boldsymbol{M}}\boldsymbol{p}_i - {}^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}\,{}^{\boldsymbol{E}}\boldsymbol{p}_i||^2}{m}} \tag{3.3} $$

In our study, three registration procedures were implemented, each one after the completion of three users. The errors recorded for these three registrations were 4.68 mm, 4.87 mm, and 4.53 mm. The registration error primarily stems from the discrepancies between the phantom and the model, attributed to manufacturing inaccuracies and the aging of the silicon phantom due to wear and tear. A secondary source of error arises from the measurement inaccuracies inherent to EM sensors. After the registration process, the virtual mesh model along with the representation of the catheter tip pose is visualized either on a conventional 2D monitor or through a HoloLens headset.

The visualizations for both standard 2D monitor and HoloLens can be seen in Fig. 3.3(d) and Fig. 3.3(e), respectively, each showcasing similar user interface designs. The user interface is divided into two sections: an external view on the left and an internal view on the right. The external view presents a comprehensive perspective of the aorta, with a thick green line illustrating the desired path for users to follow and a yellow arrow indicating the current pose of the catheter tip. The internal view offers an inside-the-vessel, forward-looking perspective akin to angioscopy. Within this view, several discrete green spheres indicate

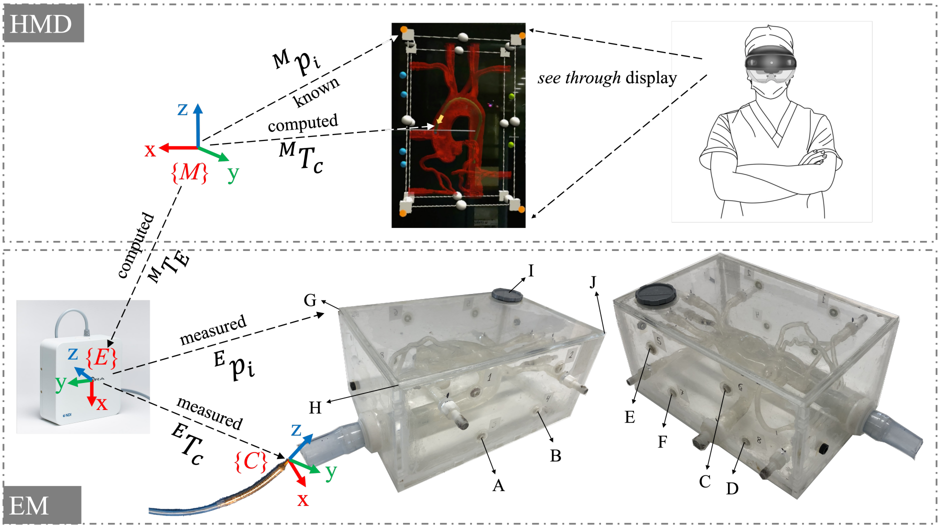


Figure 3.4: Ten markers are strategically placed on various surfaces of the phantom's outer casing, which facilitates the registration $^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}$ between the EM sensor frame $\{\boldsymbol{E}\}$ and the phantom mesh frame $\{\boldsymbol{M}\}$. By identifying the transformations $^{\boldsymbol{M}}\boldsymbol{T}_{\boldsymbol{E}}$ through registration, and $^{\boldsymbol{E}}\boldsymbol{T}_{\boldsymbol{C}}$ via EM measurements, the pose of the tip can be accurately represented in the mesh frame $\{\boldsymbol{M}\}$.

the path setpoints. Specifically, there are 26 green spheres representing the 26 setpoints derived from the path-planning algorithm. Green spheres are not uniformly distributed along the aorta. Their placement is determined by the path planning algorithm that takes into account various constraints. The path planning method is elaborated in the following paragraph. A yellow guidance arrow is added to indicate the direction and degree to which the catheter tip should bend from the current position. The white cross represents the center of the field of view. A blue sphere represents the final target at the end of the path that users must aim to reach as accurately as possible.

This path is generated through a RL method [145] to ensure the optimality of the path and the fulfillment of the constraints of the robotic catheter. The reward function of the RL algorithm takes into account several criteria: the number of steps (negative reward "-"), the number of collisions (-), whether the target position is reached (positive reward "+"), whether a path setpoint is close to the centerline (+), and whether the bending angle of the catheters is sufficient to navigate the bend of a trajectory while staying within its bending capacity (+).

It should be noted that the internal view is not present in current clinical practice. In this work, the internal view serves as an innovative feature for improving navigation capabilities within the vessels. The internal view images are captured

Table 3.2: three interactive modalities investigated in this work

	control devices	visualization devices
Mode GM	gamepad (G)	2D monitor (M)
Mode GH	gamepad (G)	HoloLens (H)
Mode HH	HoloLens (H)	HoloLens (H)

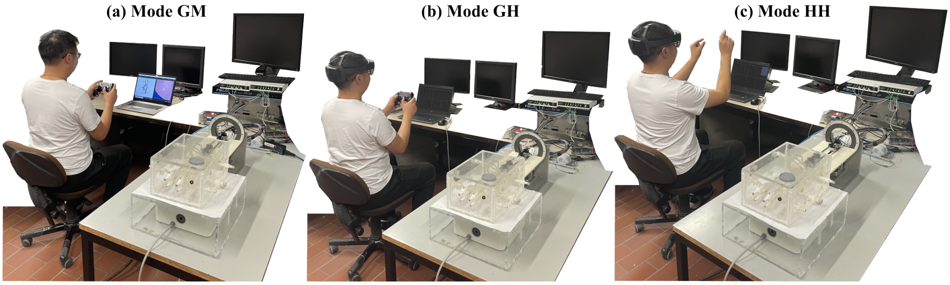

(a) Mode GM
(b) Mode GH
(c) Mode HH

Figure 3.5: Three interactive modalities investigated in this work, with their control devices and visualization devices outlined in Table 3.2.

through a virtual camera in Unity3D. The pose of the camera is continuously synchronized with the catheter tip pose, offering the operator a first-person view. In future clinical practice, the internal view could be reconstructed from IVUS or OCT images.

3.3 Experimental protocol

3.3.1 Experimental procedure

In this study, two input devices (gamepad and HoloLens) and two visualization techniques (HoloLens and 2D monitor) are investigated. The combinations of these control devices and visualization techniques could yield four possible interactive modalities. The HoloLens solely as a control device without offering visualization makes little sense and underutilizes its potential, as it is primarily considered an enhanced visualization tool, with control functionality as a secondary feature. Consequently, the remaining three interactive modalities are focused, as outlined in Table 3.2 with the corresponding experimental setup presented in Fig. 3.5. Mode GM denotes the use of a gamepad (G) as the control input device and a 2D monitor (M) for visualization. Mode GH denotes the use of a gamepad (G) as the control input device and the HoloLens (H) for visualization. Mode HH denotes the use of the HoloLens (H) both as the control input device and for visualization. The experimental procedure is as follows:

1. Participants are asked if they are interested in participating in this study and, if so, are asked to complete an informed consent form based on the standard format provided by KU Leuven. The following information is communicated to the participants: the specific tasks, namely navigating catheters in the aortic phantom, the assurance that the experiment will not involve the collection of any personal information, and their right to discontinue the experiment at any time without providing a reason.
2. The hardware, techniques, and experimental procedures are clearly explained to the participants.
3. Each participant is allocated a total of 10 minutes to familiarize himself/herself with the system and its three interactive modes.
4. Participants complete three trials for each mode, advancing to the next mode only after finishing the current one. The order of the modes varies among users: three participants adopt the sequence Mode GM-GH-HH, another three engage in Mode GH-HH-GM, while the remaining three progress through Mode HH-GM-GH. Due to practical considerations, three types of order are employed instead of six. By changing the order of the modes, the potential impact of the learning curve is distributed across all modes, making it less likely that the learning effect biases the results. In total, each user performs nine trials. The specific details of a trial are described in Section 3.3.2.
5. Upon completion of the three trials per mode, participants are asked to fill out a NASA Task Load Index (NASA-TLX) form [205]. Once the participants have completed all nine trials, they are also asked to complete a tailor-made questionnaire.

The NASA-TLX form employs a 20-point scale to subjectively evaluate the workload that users experience while performing a specific task. The evaluation comprises the following six metrics: 1) mental demand; 2) physical demand; 3) temporal demand; 4) performance; 5) effort; and 6) frustration. Participants were asked to evaluate each mode by considering the following aspects related to the six dimensions mentioned above: 1) How mentally demanding was the task? 2) How physically demanding was the task? 3) To what extent did you feel a sense of urgency or haste in the task's pacing? 4) How successful were you in completing the task? 5) How much effort did you put into achieving your performance level? 6) To what extent did one feel insecure, discouraged, irritated, stressed, or annoyed? These scores provide subjective insight into the perceived workload of users. The tailor-made questionnaire gathers the following information: the profession of a participant, whether they wear glasses, the dominant hand (left or right), prior experience with gaming and AR headsets, as well as their good/bad experiences during the experiment.

3.3.2 Detailed description of tasks per single trial

The objective of each experiment is to steer a robotic catheter within an aortic phantom from the descending aorta to the aortic root, each time with one of the three interactive modalities.

During the experiment, participants are prohibited from viewing the transparent phantom directly. Instead, they are asked to act based on visual feedback from either a standard 2D monitor or a HoloLens headset. Participants receive both external and internal visual feedback. The external view offers a comprehensive view of the aorta. A yellow arrow is used to indicate the pose of the catheter tip. The internal view provides a first-person perspective on the vessel. A predefined trajectory, depicted by a thick green line within the virtual phantom, is provided for guidance. When wearing the HoloLens, participants have the flexibility to alter their positions during the experiment to view visual feedback from various angles. The procedure of a single trial can be outlined as follows:

1. Both clamps of the catheter driver are released, and the catheter is positioned inside the catheter driver.
2. The catheter is advanced manually until its tip reaches the starting point of the planned path. To reduce friction, the catheter body is lubricated with baby oil.
3. The control system is initialized.
4. The user teleoperates the catheter. The task entails aligning the catheter tip with the predetermined trajectory until the end, following the sequence of green spheres one at a time. The user is suggested to follow the guidance provided by the yellow arrow, which indicates both the direction and the degree of catheter bending required (by the length of the yellow arrow).
5. The duration of the procedure is recorded. Users are expected to complete the experiment in three minutes. If users feel that they have not achieved the target accurately and time remains, they may decide to retract the catheter and attempt to target again.
6. The catheter is then retracted to the starting point. This retraction is executed rapidly, without meticulous control.

During insertion, improper tip movement could potentially exert an excessive force on the vessel walls. This may result in vessel rupture or dislodgement of the calcification, which is more likely to occur, particularly due to the acute tip.

3.3.3 Performance metrics

The performance metrics can be categorized into subjective and objective measures. Subjective evaluations are derived from questionnaires and the NASA-TLX form (introduced in Section 3.3.1). The objective metrics that are used to quantify the performance are:

Duration (T)

The duration of a single experiment refers to the time span from its initiation to its completion. Data recording is started and t_0 is marked when the user starts to teleoperate the robotic catheter. An experiment comes to an end when the catheter is navigated to the aortic root, and when the user confirms his/her satisfaction with the targeting accuracy through verbal confirmation. At this time, the data recording is stopped and t_g is marked. The duration is then:

$$T = t_g - t_0 \tag{3.4}$$

where t_0 and t_g are the first and last timestamp, respectively. The maximum allotted time is 3 minutes. In cases where the maximum allocated time is exceeded, the user is asked to repeat the trial.

Tracking error (T_r)

The tracking error refers to the average deviation between the actual $\boldsymbol{s_j}$ ($j = 1, ...k$ with $t = t_0$ when $j = 1$, and $t = t_g$ when $j = k$) and desired trajectory $\boldsymbol{s_i^d}$ ($i = 1, ...n$). In our experiments, the number of setpoints of the desired trajectory n is set to 500. These setpoints are interpolated from the output of the RL path-planning algorithm, originally consisting of 26 setpoints. The actual trajectory is recorded by an EM sensor at a sampling frequency f of 40 Hz. For each point $\boldsymbol{s_j}$ on the actual trajectory, the shortest distance to the desired trajectory is identified and treated as the deviation for that individual point. Tracking error is computed as the mean error of all these individual points, as illustrated in Algorithm 1.

Targeting error (T_a)

The targeting error is calculated by searching the entire trajectory and determining the shortest distance between the trajectory $\boldsymbol{s_j}$ and the target $\boldsymbol{s_n^d}$. This metric serves as a criterion to evaluate the accuracy of reaching the target.

$$T_a = \min_{\mathbf{j}} ||\boldsymbol{s_n^d} - \boldsymbol{s_j}|| \tag{3.5}$$

The significance of differences between interactive modalities is assessed through statistical tests. Because the selection of the significance test method differs based on the data distribution, the distribution of each data set was first examined to determine whether it follows a normal distribution, using the Shapiro–Wilk test

Algorithm 1: Calculate Tracking Error

```
Input: ActualTrajectory $s_j$, $j = 1, ..., k$, DesiredTrajectory $s_i^d$, $i = 1, ..., n$
Output: Tracking Error $T_r$
Initialize TotalError ← 0
for each Point $s_j$ in ActualTrajectory, $j = 1$ to $k$ do
    MinDistance ← ∞
    for each Point $s_i^d$ in DesiredTrajectory, $i = 1$ to $n$ do
        if Distance($s_i^d$, $s_j$) < MinDistance then
            MinDistance ← Distance($s_i^d$, $s_j$)
        end
    end
    TotalError ← TotalError + MinDistance
end
Result: $T_r$ ← TotalError / $k$
```

[206] with a significance level of 0.05. Subsequently, depending on their respective characteristics, we applied the Kruskal-Wallis test [207] for populations that did not follow a normal distribution, and used the t-test [208] for those that were normally distributed. A significance level of 0.05 is used for both tests.

3.4 Results and discussion of the user study

3.4.1 User profiles and subjective evaluation

Nine participants, aged 20 to 30 years and with an educational background in engineering, participated in the user study. User profiles are summarized in Fig. 3.6. Eight out of nine wear glasses, with only one participant not requiring them. Eight participants are right-handed, while only one individual is left-handed. Seven participants identified themselves as novices with no prior experience with HoloLens. One participant indicated to have an intermediate experience level, having used the HoloLens several times. One participant is highly experienced, working with HoloLens on a daily basis. As for the gaming experience, three participants are novices, four have an intermediate level of experience, and two are highly experienced. From the aforementioned statistics, it can be concluded that users have more experience with gamepads compared to HoloLens. This difference in experience may contribute to the varying levels of confidence when using these two control devices. Lastly, with regard to their experience with steerable catheters, five participants have never steered a catheter before. Two out of the nine participants possess some steerable catheter experience, and the remaining two participants interact with catheters regularly.

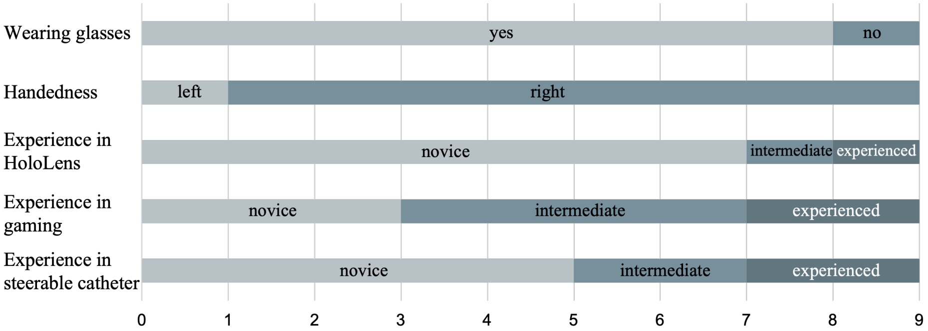


Figure 3.6: A summary of the nine user profiles includes information about their previous experience with gaming, with the HoloLens, and with handling catheters, in addition to their handedness and whether they wear glasses or not.

The perceived workload measured in NASA-TLX form is shown in Fig. 3.7. The central line illustrates the average score (20-scale) from nine users in the six different aspects of the NASA-TLX form. The colored area represents the interquartile range, with its boundaries indicating the lower and upper quartiles (i.e., 25% and 75%). It is important to note that a lower score signifies less workload perceived by users, namely better performance.

1) Evaluation from NASA-TLX: In terms of mental demand, all three modes exhibit similar performance, with Mode GM achieving the best score of 8.6 compared to Mode HH's score of 10.3. However, when considering the physical demand, Mode GM and Mode GH outperform Mode HH, suggesting that using HoloLens as an input device requires more physical effort, especially for novices. This may be due to the novices of the HoloLens continuously holding their arms up to maintain catheter control, concerning that dropping their arms might result in unrecognized hand gestures, a loss of control, and potential damage to the catheter or phantom. This observation is further supported by the ratings for physical demand associated with Mode HH. Intermediate and expert users reported a score of 5.0 ± 2.0, which differs from the novice users' score of 8.7 ± 4.0. On the contrary, gamepad users can rest their elbows at their waist and lower their hands to a more relaxed position, reducing physical strain. This trend is also evident in the grading of temporal demands. When using the HoloLens, some novice users tend to maintain a consistent catheter speed by placing speed control blocks in the same position and focus on controlling the catheter's bending, while expert and intermediate users typically possess a better knowledge of adjusting the catheter speed in response to the varying locations of the catheter tip. In contrast, all users, when using the gamepad, are more actively adjusting the insertion speed to their liking. In easier-to-navigate areas, such as the descending aorta, users tend to select a higher speed. As a result, gamepad-based catheter control is generally faster than HoloLens control. In terms of performance, user feedback provides

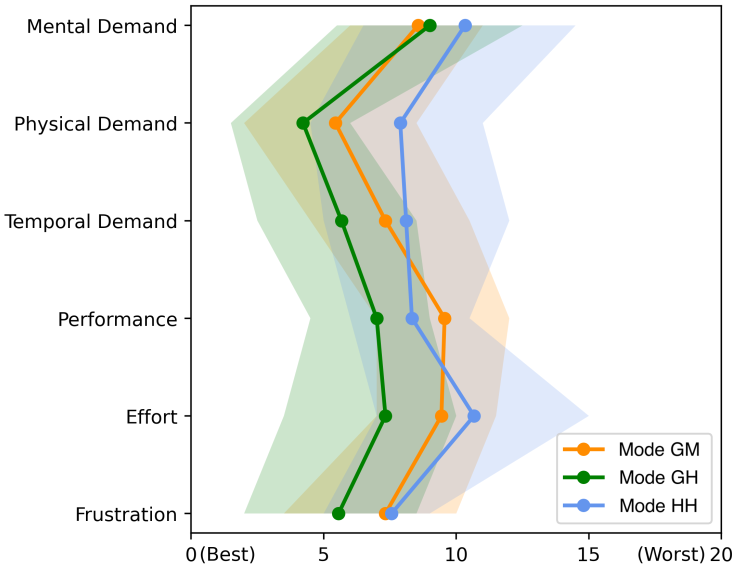


Figure 3.7: NASA-TLX scores (20-scale) represent the workload associated with three interactive modalities. The central lines signify the average scores for all nine participants in each respective mode. The colored area represents the data between the lower and upper quartile, which is known as the interquartile range. Note that lower scores indicate less workload perceived by the users.

an objective evaluation of their success in completing tasks. In this regard, the best performance is perceived as Mode GH (7.0), followed by Mode HH (8.3), and lastly Mode GM (9.6). This suggests that the 3D visual feedback from the HoloLens improves user confidence. As for effort and frustration, Mode GH remains the most favorable, with Mode GM slightly better than Mode HH. This could be attributed to the fact that most users have more experience with gamepad than with HoloLens, making catheter control via gamepad easier for them. In the scenario of HoloLens, users with limited experience may face challenges in achieving accurate hand recognition, especially when employing hand rays, a method used for interacting with holograms remotely. These issues primarily contribute to increased effort and potential frustration for users.

In general, Mode GH demonstrates the best performance in subjective evaluations among the three interactive modes with an average score of 6.5 across six aspects, with Mode GM (7.9) and Mode HH (8.8) showing comparable results. This suggests that using a gamepad for control and a HoloLens for visual feedback creates the most comfortable user experience. The standard deviations of Mode GM, GH, and HH are 3.4, 3.8, and 3.8, respectively, indicating that the difference in opinions among users is not large.

2) Results from questionnaire: In the questionnaire, users were asked to share their likes and dislikes about the experiment, providing complementary insights into their NASA-TLX scores. For control, the gamepad is generally preferred due to users' familiarity and its ergonomic design. However, the limited bending angle of the thumbstick, combined with the control strategy that maps the thumbstick's bending angle to the pressure of the artificial muscles, causes users to perceive the gamepad as overly sensitive, making fine motion control challenging. However, this may potentially be a limitation of the current implementation and is not necessarily a fundamental limitation of the interface. On the other hand, the HoloLens offers a larger panel to regulate catheter bending, resulting in higher control resolution and facilitating easier fine motion control compared to the gamepad. The users reported that, when using the HoloLens, the primary concern is the potential inaccuracy of hand gesture recognition. This may stem from users' limited experience with the device. This can cause anxiety about the possibility that the catheter may not stop and cause damage to the catheter tip or phantom due to the excessive interaction force. Despite the availability of an emergency button, users remain apprehensive as it also relies on hand gesture recognition. In terms of visual feedback, users prefer the HoloLens over a standard screen as it displays 3D trajectories and allows users to move around the holographic phantom in real-world space for more effective catheter tip alignment. In contrast, 2D screens do not offer depth perception. Although users have the option to manipulate the virtual model to observe it from different angles, this action requires an input device. In addition, this method is not as intuitive or effective as simply moving around the hologram with the HoloLens. Only one user reported dizziness from using the HoloLens. In addition, users appreciate the HoloLens feature, which allows easy relocation of holograms according to individual preferences. In our experiment, users differed in body height, as well as in their preferred placement for the holographic components. They reported that they used the relocation feature to adjust the position of the control panel and the visualization panel at their ideal location both before and during the experiment. However, there were instances where occasional modifications were made during the experiment. This customization accommodates diverse user needs and improves overall satisfaction. Conversely, for the gamepad, two users, who identified themselves as intermediate or experienced gamers, indicated a preference for an arrangement where bending control is assigned to the right thumbstick, and forward/backward movement to the left thumbstick, consistent with the design found in most computer games. Regarding the investigation of the effectiveness of the internal view, eight out of nine users agreed that the internal view was beneficial during the procedure. Only one user mostly relied on the global view instead of the internal view.

3.4.2 Objective evaluation

Figure 3.8 illustrates the experimental procedure through three distinct stages. In Fig. 3.8(a), the catheter is positioned in the descending aorta, and the internal

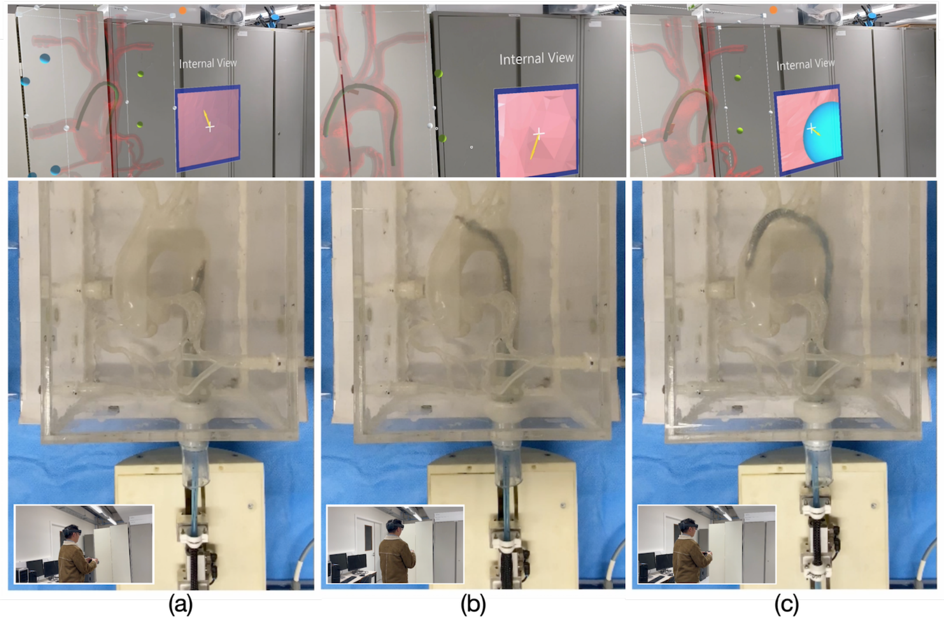


Figure 3.8: The experimental procedure with the robotic catheter at three different locations when steering using Mode GH: a) descending aorta; b) aortic arch; c) aortic root.

view features an arrow pointing to the upper left. This informs the user that the catheter must bend in that direction to follow the optimal path and navigate through the aortic arch. Figure 3.8b shows that the catheter is almost past the aortic arch, with the guidance arrow pointing towards the bottom left, indicating the need for increased bending. In Fig. 3.8(c), the catheter successfully reaches the aortic root, as demonstrated by the corresponding HoloLens view where the blue target is nearly achieved. An arrow pointing towards the bottom-right direction suggests that by gently bending the catheter in this direction, the target can be reached.

Data collected during the experiment were analyzed to obtain performance metrics, as detailed in Section 3.3.3. Figure 3.9 illustrates the tracking error, the target error, and the duration for three interactive modes, achieved by the nine participants. Regarding the tracking error, Mode GH displayed superior performance among all three interactive modes, with a median of 4.92 mm, followed by Mode HH (5.17 mm) and Mode GM (5.39 mm). However, Mode GH also demonstrated the highest variance among the three modes, indicating considerable tracking performance disparities between users or trials, while Mode HH exhibited the lowest variance. These observations are in line with the results depicted in Fig. 3.10. The full trajectory captured by the EM sensor for a single trial is shown in Fig. 3.10,

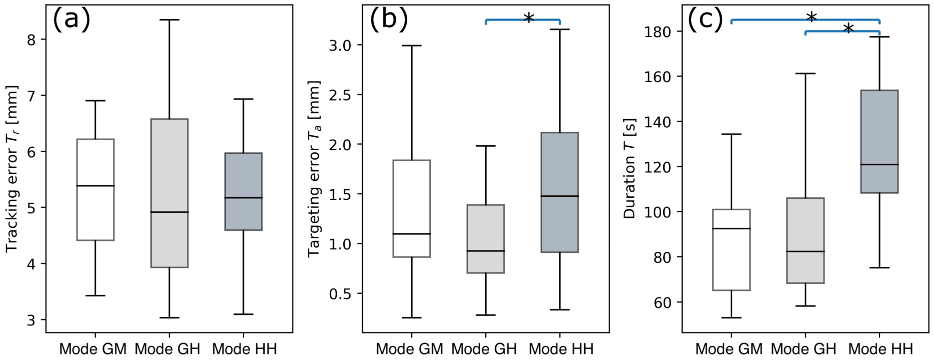


Figure 3.9: The tracking error, targeting error, and duration of the three interactive modalities are presented in a box plot, which displays the median, interquartile range, and minimum-maximum range of the data. If two groups are connected with a line and an asterisk, this indicates a significant difference between them. (i.e. $p < 0.05$ using Kruskal-Wallis test or t-test)

with subfigures (a) to (c) representing Mode GM, GH, and HH, respectively. Figure 3.10(c) (Mode HH) reveals a smoother trajectory compared to Fig. 3.10(a), (f) (Mode GM). This can be explained as follows: both control strategies (gamepad vs. HoloLens) establish a proportional relationship between the commands generated by the control input device and the pressure applied to the artificial muscle. In the HoloLens control, the green sphere for catheter control remains stationary upon release, resulting in a generally smoother trajectory. On the contrary, the gamepad exhibits a spring effect, automatically returning to its neutral position when released. Consequently, improper control may produce a zigzag-shaped trajectory, as demonstrated in Fig. 3.10(f). However, as depicted in Fig. 3.10(a), after multiple trials, the gamepad control can also achieve a smooth trajectory and performance comparable to that of the HoloLens-based control. From another point of view, the gamepad control surpasses HoloLens in terms of responsiveness and reliability. Maintaining consistent hand gesture recognition using HoloLens was found to be difficult for novices. This could be explained by the user feedback and the observations made during the user study: 1) users may not be familiar with the interacting principles of HoloLens, namely may not interact with holograms as if they were tangible objects. Consequently, when the hologram is not within the user's physically reachable range (i.e., limited by the length of arms), users sometimes still attempt to interact directly with the hologram; 2) users may not be fully accustomed to the HoloLens hand gesture command known as "air tap", which allows the user to pinch an object using the thumb and forefinger. However, during the study, some users were unable to perform such a motion, thus impeding their interaction with the hologram. On the other hand, the thumbsticks and buttons of the gamepad control consistently provide a responsive

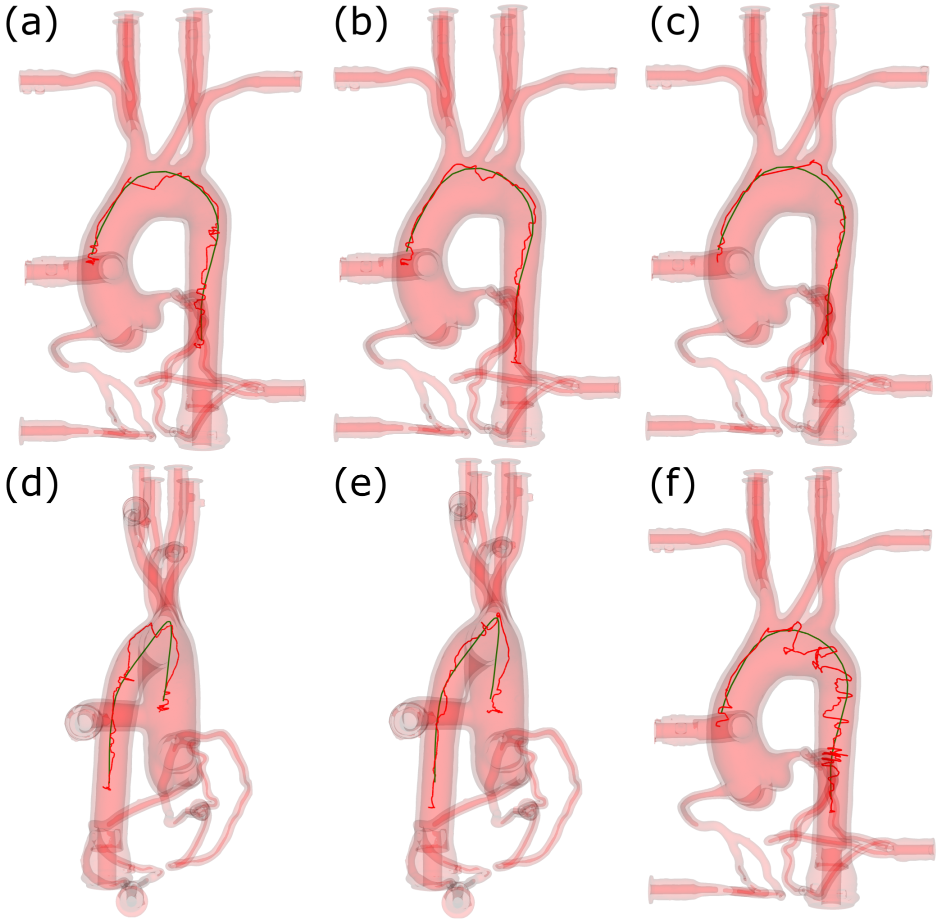


Figure 3.10: The recorded trajectories of a single trial are presented, with subfigures (a), (b), and (c) corresponding to Modes GM, GH, and HH, respectively; (d) side view of a; (e) side view of b; (f) another example in Mode GM with a zigzag-shaped trajectory due to improper control with gamepad.

interaction. Furthermore, it can be observed that the HoloLens significantly aids in reducing the tracking error, particularly in the depth direction (Fig. 3.10(d-e)). This can be substantiated by the tracking error in the depth direction, which was additionally calculated across the three interactive modes. Mode GH and Mode HH exhibited comparable performance, with median errors of 2.55 mm and 2.41 mm, respectively. In contrast, Mode GM performed worse, registering an error of 2.81 mm. Figure 3.10(d) and Fig. 3.10(e) display the side view of Fig. 3.10(a) (Mode GM) and Fig. 3.10(b) (Mode GH), respectively, with visual feedback provided by a standard 2D monitor and HoloLens. One can observe that their performance in

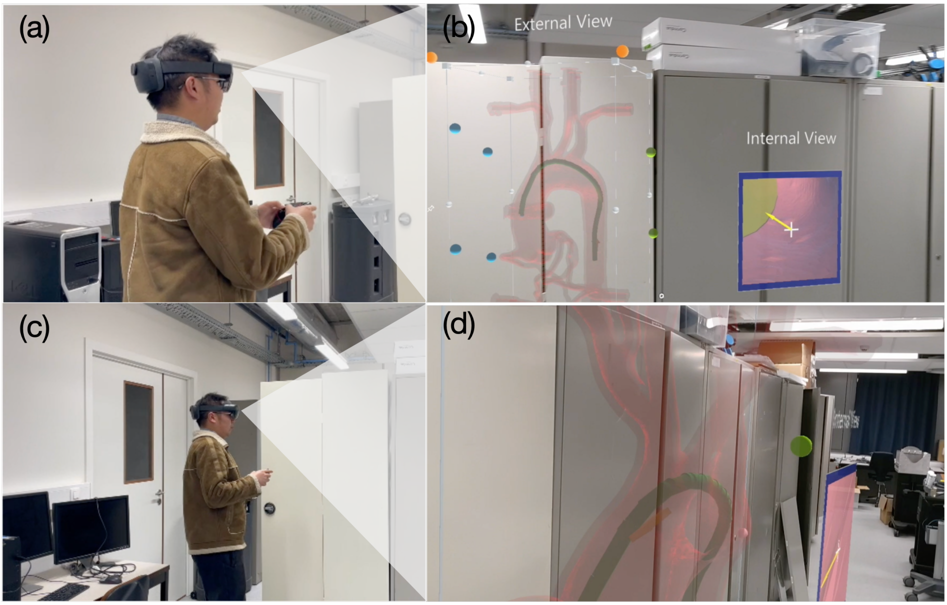


Figure 3.11: The HoloLens provides users with the ability to view visual feedback from various angles: (a)-(b) Users can view the holographic phantom from a frontal perspective. Although it maintains the same perspective as seen on a 2D monitor, the HoloLens offers a three-dimensional viewpoint. (c)-(d) By moving in real space and viewing the holographic phantom from various perspectives, users gain the ability to align the catheter tip along the trajectory from all angles, rather than merely aligning it from a frontal view.

the side view varies considerably. With HoloLens, the user effectively aligns the tip with the predefined trajectory in the side view. Conversely, on a 2D screen, the user may deviate from the trajectory in the side view, even if they believe that they have achieved satisfactory performance by solely confirming the frontal view. The results substantiate the benefits of HoloLens feedback, primarily when users move around the holographic phantom in real space, thereby viewing it from various angles. This user behavior is illustrated in Fig. 3.11. When paired with a gamepad, which offers high responsiveness and can be easily carried, optimal performance was achieved. On the other hand, this also suggests that there is room for improvement in the quality of 3D visualization; otherwise, users would not need to adjust their perspective by moving. In our future research, we plan to implement an effective method for rotating the view displayed on the 2D screen and subsequently draw a comparison with the 3D visualization.

Regarding the targeting error, Mode GH exhibited the lowest error at 0.93 mm. Mode GM, with an error of 1.10 mm, outperformed Mode HH (1.48 mm). The

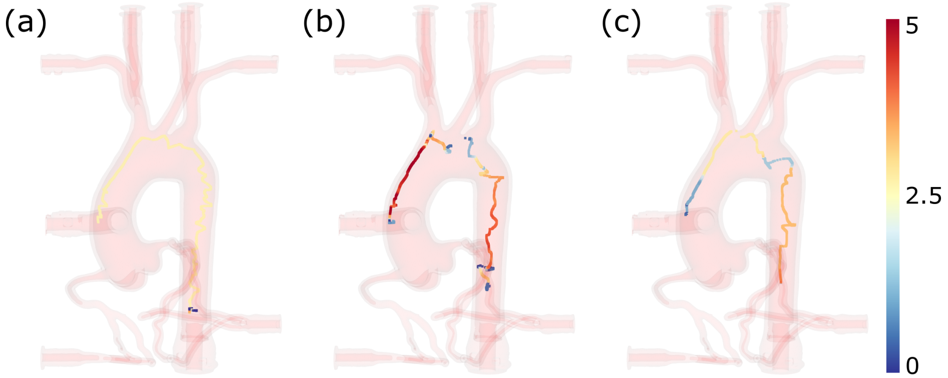


Figure 3.12: The catheter insertion speeds (0 - 5 mm/s) in Mode HH, as performed by users with varying levels of experience with HoloLens, are depicted as follows: (a) novice user, (b) intermediate user, and (c) expert user.

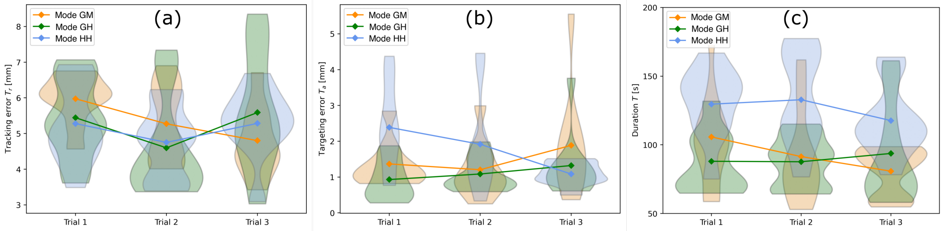


Figure 3.13: The user performance, as a function of the number of trials, depicted using violin and line plots for comprehensive visualization.

significance test demonstrated a statistically significant difference between Mode GH and Mode HH. The findings indicate that the gamepad demonstrates superior targeting ability compared to the HoloLens. When approaching the target, users can easily stop the catheter driver's insertion, allowing ample time to modify the catheter tip's pose. Once prepared, they can advance the catheter driver, achieving the target rapidly. On the contrary, the majority of users find that stopping the driver and adjusting the orientation of the tip require synchronized movements of both the left and right hands. This involves handling hand gesture recognition simultaneously, which could be challenging given their limited familiarity with the HoloLens. As a result, they often attempt to target while the motion of catheter insertion is still ongoing, leading to increased targeting errors in Mode HH.

Concerning the duration, Mode GH (82.34 s) and Mode GM (92.45 s) demonstrated similar performance levels, with both being statistically distinct from Mode HH (120.87 s). This can be attributed to the ease of controlling the driver's speed using a thumbstick, as users can simply bend the thumbstick to adjust the speed. Stopping the driver is also straightforward, as users only need to release the thumbstick. On the other hand, when operating the system using the HoloLens, users frequently encounter challenges in dragging the slider back to the neutral region, where the speed is set to zero. Figure 3.12 illustrates the varying catheter insertion speeds of users with different levels of experience in HoloLens, arranged from left to right as novice, intermediate, and expert users. The color heat map represents speed, with red indicating high speed, yellow indicating medium speed, and blue symbolizing low speed. One can observe from the figure that a user's level of experience impacts their steering style. Part of the novice users adhere to a consistent and moderate speed of about 2.5 mm/s, directing their primary attention towards the control of bending maneuvers. Intermediate and expert users demonstrate the ability to adjust their speeds flexibly according to different regions.

Figure 3.13 illustrates the performance of the users as a function of the number of trials to determine the presence of any learning curves. Figure 3.13 features a combination of a violin plot [209] and a line plot. The line plot represents the average value, while the area of the violin plot illustrates the data distribution. In terms of targeting error, Mode HH exhibits the most pronounced learning curve. In addition, the data distribution area narrows as the number of trials increases, indicating fewer outliers. In the third trial of Mode HH, all the trials have an error of approximately 1 mm. The experimental duration demonstrates the most significant learning curve. For both Mode GM and Mode HH, the duration decreases considerably as the number of trials increases, as evidenced by both the line plot and the data distribution. In the case of Mode GH, although the average duration in the third trial increases, the data distribution reveals that this is due to outliers substantially raising the average value. Most users complete their experiment in around 100 s. The maximum duration of 164 s in trial 3 is because the user was dissatisfied with his/her initial target performance and was prompted to retract the catheter and attempted another trial. In general, it is quite challenging to analyze the learning effects through three trials. In future work, we will attempt to conduct more experiments to better analyze learning effects.

In summary, the findings of this chapter can be outlined in the following bullet points. Note that these findings hold within the limitation of this study, which was conducted with a limited number of participants.

- the combination of a gamepad and a HoloLens demonstrated the highest popularity and performance;
- the gamepad surpasses the HoloLens, owing to its familiarity from gaming, exceptional responsiveness (compared to hand gesture recognition of the HoloLens), and ergonomic design that enables users to maintain comfortable gestures;

- the HoloLens shows superior performance in control by offering high-resolution control for fine motions and the freedom to reposition different holographic control components on the spot;
- the HoloLens significantly surpasses the 2D monitor by providing: a) overlaid 3D images of the patient's anatomy and catheter position within the physician's field of view, creating a 360-degree interactive environment with depth perception, enabling multi-angle examination of holographic objects, and b) hands-free interaction with holographic models, which removes the reliance on a mouse, a keyboard or a gamepad, as is necessary with a 2D monitor.

3.5 Conclusion on comparative analysis of interactive modalities for intuitive endovascular interventions

Fluoroscopy has been crucial for catheterization procedures, yet it exhibits certain limitations. First, it is unable to provide depth perception. Second, fluoroscopy is associated with ionizing radiation. To address these challenges, employing teleoperation for catheter steering and using 3D visualization devices could provide substantial benefits, while there are some issues such as the limited control resolution of the employed gamepad. The influence of different types of visual cues was also not investigated.

In this chapter, three implementations that use different control and visualization devices are compared through a user study. An endovascular catheterization system, comprising a robotic catheter, HoloLens, and a gamepad, is constructed for experimental validation. The user study revealed that the greatest appreciation and performance are achieved for the combination of a gamepad as the control device and HoloLens as the visualization device. By moving within the physical space, users can explore various perspectives of holographic imaging through the HoloLens. The gamepad, on the other hand, is recognized as an easy-to-use and intuitive control device that is highly responsive and portable.

The findings of this user study have the potential to improve current endovascular intervention procedures by introducing an innovative clinical workflow. In this approach, physicians remotely maneuver the catheter using a gamepad while wearing a HoloLens. Within HoloLens, a holographic representation of the anatomy is superimposed onto the patient's body. The catheter's configuration, such as tip position and shape, can be acquired through nonradiative methods like EM tracking or FBG sensing and subsequently represented in the holographic anatomy. Guided by AR, physicians could potentially navigate catheters and guidewires using an intuitive teleoperation approach.

Validation of the impact of AR and teleoperation in an *in-vitro* vessel navigation study remains difficult owing to the limited availability of robust steerable catheter systems, the recent availability of high quality AR-HMD as well as the complexity of integrating these components. Despite these challenges, we have integrated these components and devised an AR interface, which results in a catheterization system capable of operating in various interactive modes. In this study, the guidance cues are uniformly implemented across both 2D and 3D visualization methods. Therefore, the comparison purely lies in the distinctive characteristics of 2D and 3D visualization. For a more comprehensive understanding, additional studies are required on different types of guidance cues. Another limitation of the work presented in this chapter is the limited number of studies conducted so far. Given the delicate and complex structure of vessels, intuitive catheter steering becomes crucial in these applications. Therefore, it would be valuable to conduct an increased number of trials per user and engage a broader range of participants, including clinicians, in future research. Moreover, the limited range of motion of the thumbstick may have led to the zig-zag catheter trajectory observed in our user study, which needs further investigation. Initially, using control input devices with a larger range of motion should be explored. Subsequently, the second investigation should focus on using a single component for catheter control, such as the Virtuose 6D robot (Haption Inc., France) or the Touch haptic device (3D Systems, USA) instead of controlling catheter bending and insertion via two thumbsticks. Additionally, it would be advantageous to extend this investigation to *in-vivo* animal experiments.

Although this chapter contributed to sub-objective 2, the visual feedback to users is still limited by displaying only the tip pose of the catheter, represented by an arrow. While this provides some useful information, it could make physicians lack the comprehensive awareness that could be achieved by displaying the catheter's entire shape. Offering a full visualization of the catheter's shape could significantly enhance the procedure by improving accuracy, reducing the risk of tissue damage, and facilitating complex maneuvers. Additionally, it could speed up the procedure by assisting in quick decision-making, and increase clinician confidence. Therefore, the next chapter should aim to provide visual feedback that encompasses the catheter's entire shape, ideally utilizing a non-radiative method. This advancement would not only enhance the procedure, but would also align with the growing emphasis on minimizing radiation exposure for both clinicians and patients.

Chapter 4

Improved proprioceptive sensing of catheters using deep learning

This chapter is based on the following publication:

X. T. Ha*, **D. Wu***, M. Ourak, G. Borghesan, J. Dankelman, A. Menciassi, E. Vander Poorten, “Shape Sensing of Flexible Robots based on Deep Learning”, *IEEE Transactions on Robotics*, vol. 39, no. 2, pp. 1580-1593, 2022

In this chapter, a deep learning method for shape sensing of continuum robots based on multi-core Fiber Bragg Grating (FBG) fiber is introduced. The proposed method, based on an Artificial Neural Network (ANN), differs from traditional approaches, where accurate shape reconstruction requires a tedious characterization of many characteristic parameters. A further limitation of traditional approaches is that they either require multiple fibers, whose location relative to the centerline must be precisely known (calibrated) or a single multi-core fiber whose position typically coincides with the neutral line. The proposed method addresses this limitation and thus allows shape sensing based on a single multi-core fiber placed off-center. This helps in miniaturizing the robot and leaves the central channel of the robot available for other purposes. The proposed approach was compared with a recent state-of-the-art model-based shape sensing approach. A 2-DOF bench-top fluidics-driven catheter system was built to validate the proposed ANN. The proposed ANN-based shape sensing approach was evaluated on a 40 mm long steerable continuum robot in both 3D free-space and 2D constrained environments, yielding an average shape sensing error of 0.24 mm and 0.49 mm, respectively. With these results, the superiority of the proposed approach compared to the recent model-based shape sensing method was demonstrated.

This chapter is in alignment with Sub-objective 3 of the thesis, which focuses on reconstructing the 3D shape of catheters using non-radiative methods. This approach could reduces radiation exposure for both clinicians and patients, thus enhancing the safety of interventions. Moreover, particularly in complex environments such as vasculatures, possessing real-time knowledge of the catheter's 3D shape and motion provides essential feedback for physicians, crucial for effective catheter control.

4.1 Introduction

In applications where flexible endoscopes or catheters are used, surgeons or interventionalists require a good understanding of the input-output behavior of these robots. Furthermore, interventionalists wish to know in detail which space is occupied by the robot e.g. to judge the vicinity to critical structures or to plan adequate motion commands. Real-time knowledge of the robot's shape would greatly facilitate this task. Different methods have been proposed in the literature to model the kinematic and dynamic behavior of soft and continuum robots. However, accurate models are complex and computationally intense [210], which limits their intraoperative use. Moreover, models are designed for a specific instrument and may require dedicated modeling efforts to characterize the input-output behavior and the distributed shape for each individual instrument.

Different sensing modalities based on imaging (e.g., fluoroscopy) [211], electromagnetic (EM) tracking system [212], electrical impedance tomography (EIT) [213] or bending resistance [214] have been proposed to acquire the shape of interventional devices. Unfortunately, each sensor modality suffers from its own drawbacks. Fluoroscopic systems have the downside of emitting harmful radiation and providing only a 2D view of the patient/instrument. A Computed Tomography (CT) scan is even more harmful. EM systems do not offer uniform accuracy and are sensitive to the presence of metallic material. These problems complicate adoption in clinical practice. The use of EM and EIT for shape sensing requires different coils or electrodes to be integrated into the structure. This can be cumbersome and may increase the size, while also complicating the design of the robotic system. Gerboni *et al.* [214] integrated commercial flex bending sensors in their soft robot. While these sensors are appealing, they exhibit drift and hysteresis, which requires additional modeling and identification efforts to precisely compute the bending.

In recent years, optical fibers with Fiber Bragg Gratings (FBGs) have become popular for shape sensing. Properties that make FBG optical fibers particularly attractive for integration in continuum robots are their bio-compatibility, compact size, light weight, flexible nature, and safety, e.g. being free from the risk of electrocution. In addition to real-time knowledge of the shape, FBG fibers can also be used advantageously to characterize and control the behavior of soft and continuum robots. However, FBGs suffer from spectral distortion when bending

[215]. Another major challenge of FBG-based shape sensing is that strain sensing through traditional FBGs cannot distinguish strain caused by twist from strain caused by bending. Mode *et al.*[216] proposed a method to sense the 3D shape of a continuum robot that incorporates bending, twisting, and elongation. The method proposed by Mode *et al.* is, however, only verified by simulation. Further limitations to the overall estimation accuracy are due to uncertainty in sensor location during assembly.

A traditional FBG-based shape sensing method starts with measuring the wavelength shift and follows by calculating the strains at discrete locations where FBGs are inscribed. Finally, a set of curvatures and angles of the bending plane at these locations is estimated. The final shape of the robot is reconstructed by integrating curvatures and bending plane variation over the fiber length. In these traditional methods, several characteristic parameters need to be identified carefully, such as: strain sensitive coefficients, the different distances of the outer cores to the central core, the angle of each core with respect to a reference axis (referred to as x-axis in this work) and the location of the grating. Typically, some values are provided by the optical fiber's manufacturer. However, there are three elements to consider: *(1)* these parameters may vary due to the assembly process and thus require a manual calibration. Also, as demonstrated in [217], for good performance, one may have to tackle the problem of "intrinsic twist". The twist may be induced during assembly when the fiber is integrated in a flexible instrument. Also this calls for a manual calibration as deviations from the true value will have an immediate and significant impact on the shape sensing performance. *(2)* Further, to sense the shape of the continuum robot, the optical fiber is normally inserted into the center channel of the continuum robot to ensure that the shape of the optical fiber follows the shape of the robot. As a result, the center channel is occupied and hence unavailable for other purposes. For some catheter designs, such as non-assembly 3D printed surgical devices [218], in which a thin central beam serves as the continuous backbone of the steering segment, it is infeasible to place a shape-sensing fiber centrally. In such a case, adding a central channel would excessively increase the bending stiffness of the design. Due to the aforementioned reasons, methods where the constraint to place the shape sensing fiber centrally is not needed and where the precise location of the fiber, with respect to the centerline, could be characterized in a more automatic fashion would be highly advantageous. *(3)* Liu *et al.* [219] and Farvardin *et al.* [220] used a traditional FBG-based shape reconstruction method to estimate the shape of a 35 mm long continuum robot. In both works, when the continuum robot moves in a constrained environment, an increase in the distal tip tracking error is observed. However, the shape reconstruction capability should keep its quality independent of whether the continuum establishes contact with the environment.

Recently, deep learning, which is usually implemented using Artificial Neural Networks (ANNs), has become increasingly popular. Thanks to the ability of ANNs to model complex nonlinear behavior, good prediction accuracy has been reported in [221]. Another advantage of deep learning compared to traditional learning techniques is that it does not require feature extraction processes. Due to its advantages, Sefati *et al.* proposed a data-driven shape sensing method based on FBGs to solve some of the problems listed above [44]. Sefati *et al.* modeled a 1-DOF continuum robot introduced in [222, 223] as a pseudo-rigid body. The trained model was capable of estimating the distal end position of the robot. Afterwards, the robot shape is reconstructed from the distal end estimation by solving the constraint optimization for the joint angles. A major obstacle to extending Sefati's technique to multi-DOF robots is the computationally intensive optimization stage it requires.

To address the previously mentioned challenges, a novel data-driven method is introduced to predict the 3D shape of a 2-DOF continuum robot using measured wavelength shifts from the FBGs. What differentiates this approach is its ability to gather data and undergo training while the robot operates in free space. Not only is this method efficient and straightforward, but as demonstrated in this study, it also effectively adapts to more complex contact situations. The primary contributions of this chapter are:

- a new data-driven ANN-based FBG auto-calibration method is introduced that estimates the curvature and the angle of the bending plane directly from measured FBG wavelength shifts, regardless of the fiber's characteristic parameters. The proposed approach also makes it possible to get accurate shape sensing capability with a single off-centered multi-core fiber, supporting miniaturization and allowing freeing up central channels for other uses.
- experiments that compare and validate both methods are conducted, the transferability of the ANN-based shape sensing method to contact situations is demonstrated. The ANN is trained in a free-space scenario, hence requiring limited training time/data, the model shows equally good shape sensing capability when the robot contacts an unknown environment.
- ablation studies were carried out to enhance the interpretability of the deployed ANN.

This chapter is organized as follows: Section 4.2 details the design of a 2-DOF continuum robot and the experimental setup that was built to evaluate the proposed method. Section 4.3 describes the method for collecting training and validation data. Next, the proposed auto-calibration method based on a neural network is explained in Section 4.4. Experimental results are shown in Section 4.5 that compares the shape sensing accuracy of the two described approaches. Section 4.6 concludes the chapter and presents some directions for future work.

4.2 Experimental setup

4.2.1 Purpose-built steerable 2-DOF catheter segment

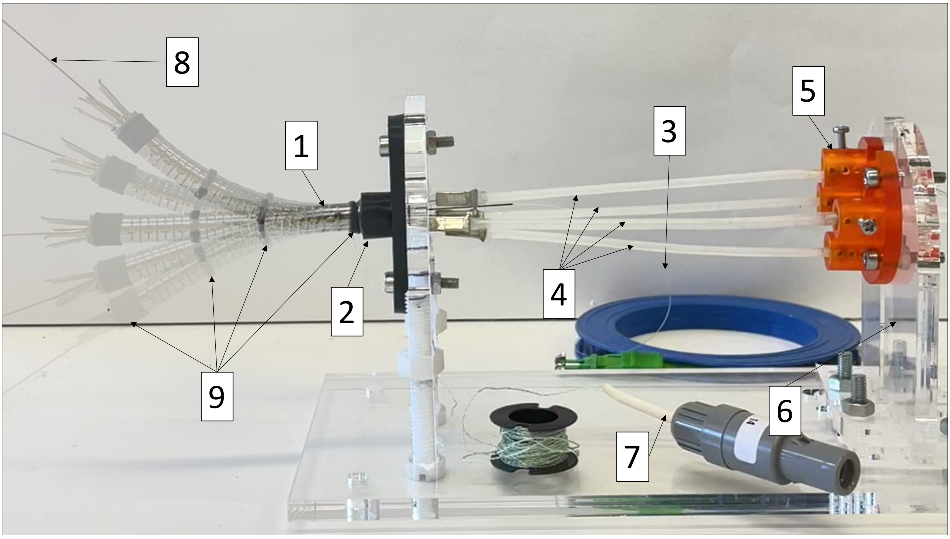


Figure 4.1: A 2-DOF pneuamtically-driven catheter system for data collection and validation: 1. Nitinol distal catheter segment; 2. catheter segment clamp; 3. FBG fiber; 4. Pneumatic Artificial Muscles (PAMs); 5. PAM clamp; 6. adjustable backplate for fine-tuning the pre-tension of the cable connected to the catheter tip; 7. EM sensor connector; 8. protective nitinol tube to accommodate the FBG sensors; 9. heat shrink to fix the protective Nitinol tube.

To collect data for training and testing auto-calibration techniques, a dedicated 2-DOF bench-top experimental setup was built (see Fig. 4.1). Without loss of generality, the setup features only the distal steerable segment of the envisioned active catheter. The 2-DOF of the catheter's distal segment are actuated with four embedded Pneumatic Artificial Muscles (PAMs). The catheter is fabricated out of Nitinol using metal laser cutting technology. Notches are made on the Nitinol segment so that the distal portion becomes bendable while minimally affecting the compression stiffness. Since this chapter aims at efficiently solving the problem of shape sensing, the design of the bendable Nitinol segment is not described here. However, details on the design can be found in [200]. The length of the flexible part (where the notches were cut) and the diameter of the catheter are 40 mm and 6 mm, respectively. With a steerable length of 40 mm, applications in minimally invasive orthopedic [44] or ENT [224] could be envisioned. Four custom-made PAMs, which are made up of a bladder, an uninflatable tube, braids, and ferrules, are used to actuate the catheter. The produced PAMs have an initial diameter

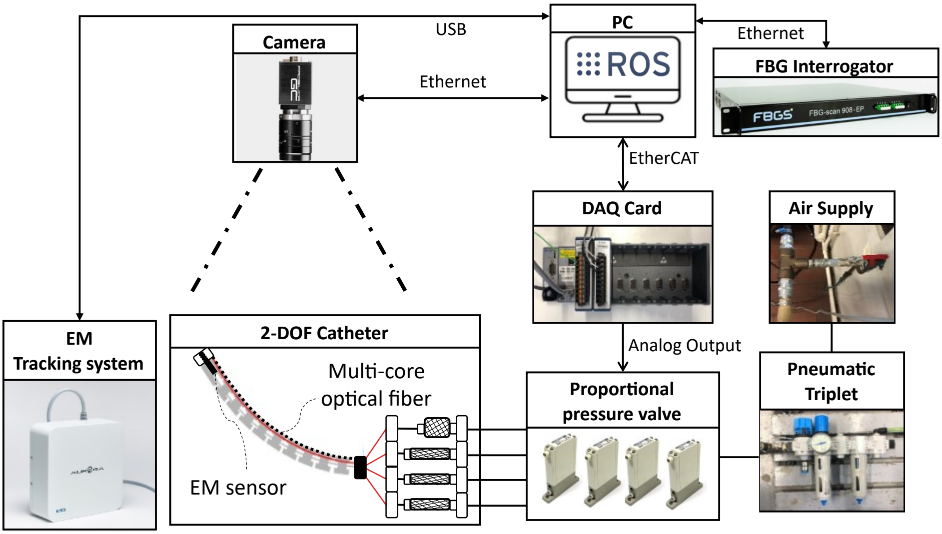


Figure 4.2: Control and actuation system for the 2-DOF catheter system: A 2-DOF PAM-driven catheter segment is controlled by a set of 4 proportional pneumatic valves, which receive command signals from a PC through a NI® CompactRIO system. The PAMs are attached off-center via four steel cables to the catheter tip. When contracting, they apply a bending moment to the tip. The resulting catheter tip pose is captured by an EM sensor. The FBG data and images of catheter are gathered by an FBG interrogator and a camera, respectively. A PC running ROS is used to gather sensor data.

of 2.1 mm and a length of 129 mm in the non-pressurized condition. The PAMs are connected off-center to the tip of the catheter by steel wires. When pressure increases, the PAM contracts and thus applies a bending moment on the catheter tip. The other end of the PAM is fixed to a movable plate, so that the pre-tension on the PAMs can be adjusted. The four PAMs are configured at an angle of 90 degrees in two antagonistic pairs, so that a 2-DOF spatial bending can be implemented. To accurately control the input pressure, the pressure is fed by an air supply, through a pneumatic triplet (Festo Corporation, Germany) as shown in Fig. 4.2. The proportional valve receives control signals from an analog output module NI-9263 (NI, Texas, USA). An EM tracking sensor (6-DOF, 0.8 mm× 9 mm, Northern Digital Inc., Ontario, Canada) is fixed to the center channel of the catheter distal tip to measure the catheter tip pose.

4.2.2 Experimental platform

To generate ground truth data for training a neural network that encodes the shape of the steerable 2-DOF catheter, an experimental platform was built as shown in Fig. 4.2. The setup consists of 3 modules: an interrogator (FBG-Scan 804, FBGS, Geel, Belgium) to measure the wavelength of the distributed FBG sensors, an EM tracking system and a 2-DOF distal catheter system. To showcase that there is no need to position a shape sensing sensor centrally, a 250 micrometer four-core fiber consisting of 20 FBGs from FBGS Inc. (Geel, Belgium) is placed outside the catheter. The fiber is fixed to the catheter by means of heat shrink tubes (Fig. 4.2). When the catheter was at rest, the fiber was parallel to the neutral axis. During the data collection process, the FBG data was measured by the interrogator at a frequency of 200 Hz while the EM sensor poses were recorded at 40 Hz. The images from the camera were streamed at 30 Hz. The recorded data was time-stamped and pre-processed to synchronize the FBG data with the corresponding EM sample using MATLAB (The MathWorks, Inc., Massachusetts, United States). All experiments are carried out on a computer with an Intel 2.1 GHz core i7 processor and 32 GB of RAM, running Ubuntu 16.04. Each core contains five gratings distributed 10 mm apart from each other. Note that each of the 20 FBG gratings has a different center wavelength. A planar NDI Aurora EM system (Northern Digital Inc., Ontario, Canada) was positioned below the steerable catheter and a monocular Prosilica camera (Allied Vision Technology, Germany) was placed approximately 30 cm above the setup. The camera faces downwards to capture the 2D shape of the catheter in the camera's image plane.

4.3 Model-based off-center fiber shape reconstruction

A new approach is proposed in Section 4.4 that takes the wavelength shifts of four gratings at one given arc-length or cross-section as input for an ANN and directly estimates the catheter's curvature and angle of the bending plane, at this arc length. The new approach is compared with an extension of a more traditional model-based approach, which is introduced in this section. In this work, to speed up the training process, we exploit the property that during free-space motion the catheter tends to follow a constant-curvature shape [225]. An experiment, described in Section 4.3.1, validates this constant-curvature assumption for free-space motion. From the constant curvature property, a method to estimate the shape and the angle of the bending plane from the pose of the EM sensor attached at the tip is then presented in Section 4.3.2. In fact, due to the constant curvature assumption, the measured curvature and the angle of the bending plane should in principle be the same at each time step for all gratings, at least during the generation of training data. This knowledge is then used to speed up training. The same constant-curvature model

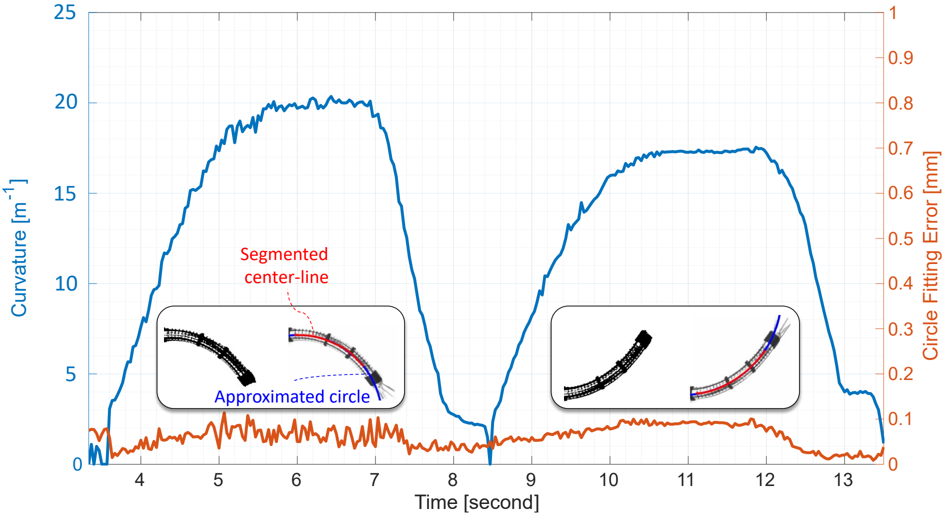


Figure 4.3: Result of constant curvature verification experiment. The threshold images and the fitted circles can be seen in the inserts. The estimated curvature and errors of circle approximations are shown respectively in blue and red.

will also be exploited as ground truth when validating shape-sensing in free-space movement. Data collection for training and validation is described in Section 4.3.3.

Note that, because the multi-core fiber is not attached to the catheter's center channel, the computed set of curvatures will not reflect the real catheter curvature and hence cannot be straightforwardly used in traditional model-based shape reconstruction approaches. An additional calibration approach is presented in Section 4.3.4 that estimates the pose of each grating relative to the catheter's central line. Based upon this information the curvature information can then be mapped to the center line. From this, the 3D shape reconstruction method, described in Section 4.3.5, can be derived. Also note that while the proposed implementation can be considered a logic extension for anyone crafted in the art, of conventional shape reconstruction, the authors are not aware of earlier attempts to implement such a model-based approach with an off-center multi-core fiber. Details on the developed ANN are given in Section 4.4.

4.3.1 Constant curvature model verification

To verify that the catheter follows the constant curvature model when operating in free space, the catheter was bent in a plane that is parallel to the image plane of the monocular camera. The camera was calibrated before the experiments to obtain the intrinsic and extrinsic parameters of the camera using MATLAB's camera calibration toolbox [226]. The captured images are first undistorted and processed

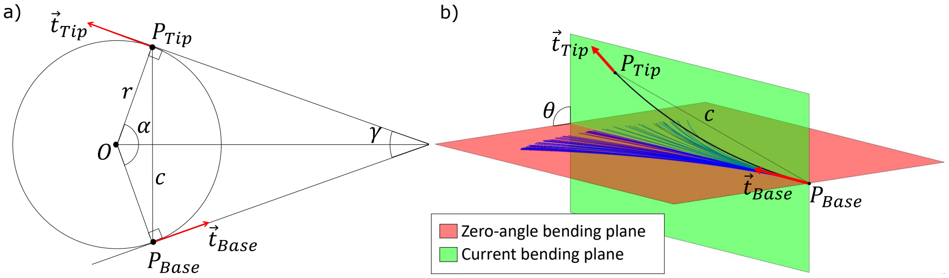
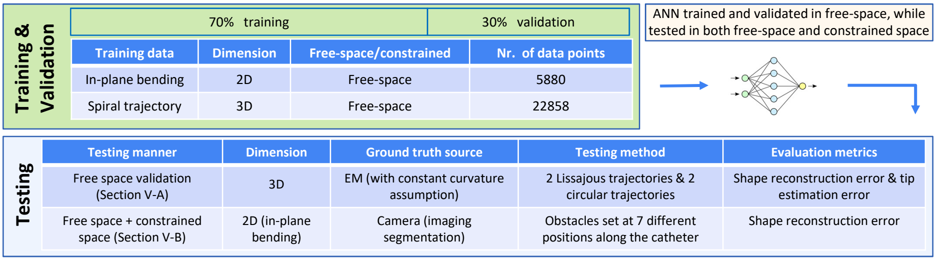


Figure 4.4: The calculation of the curvature and the angle of the bending plane, which are explained in Section 4.3.2: (a) curvature calculation; (b) θ is the angle between the zero-angle bending plane and the current bending plane.

to extract the catheter shape in 2D by color segmentation. Each image contour is evaluated to find the contour that encloses the largest image area. The largest contour is then considered the true contour of the catheter. The contour is then skeletonized by the algorithm proposed by Lee *et al.* [227] to obtain the backbone of the catheter. Since in this experiment the catheter moves in a plane that is parallel to the image plane, the 3D shape is reconstructed with a constant and known camera depth distance. The obtained set of points are then fitted with a constant curvature arc. During the experiment, different pressure values from 0 to 3.5 Bar were applied to the PAMs and the circle fitting errors were recorded. The experimental results are visualized in Fig. 4.3. The average circle fitting error was calculated by averaging the distance between the set of points belonging to the skeleton and the approximated circle. Average circle fitting errors of 0.061 mm with a standard deviation of 0.025 mm were recorded during the constant curvature model verification experiment. This low circle fitting error supports the argument to consider the free space shape as a constant curvature arc.

4.3.2 Shape and bending plane angle estimation from EM

During the ground truth generation process, the catheter is bent in free space, so the constant curvature model is valid. From the constant curvature assumption one expects that the measured curvature and angle of the bending plane of the ground truth data are the same for all gratings. From the EM sensor, attached at the tip, the catheter tip pose can be measured. Knowing the length of the catheter and measuring the tip pose when the catheter is positioned in a straight configuration, one can calculate also the pose of the base of the steerable catheter segment. Note that the base was conveniently kept fixed during the experiment. Since pressure values only affect the steerable portion, while other parts of the catheter were fixed, the base pose did not change during the ground truth generation process.



ANN trained and validated in free-space, while tested in both free-space and constrained space

Training data	Dimension	Free-space/constrained	Nr. of data points
In-plane bending	2D	Free-space	5880
Spiral trajectory	3D	Free-space	22858

Testing manner	Dimension	Ground truth source	Testing method	Evaluation metrics
Free space validation (Section V-A)	3D	EM (with constant curvature assumption)	2 Lissajous trajectories & 2 circular trajectories	Shape reconstruction error & tip estimation error
Free space + constrained space (Section V-B)	2D (in-plane bending)	Camera (imaging segmentation)	Obstacles set at 7 different positions along the catheter	Shape reconstruction error

Figure 4.5: The training, validation, and testing procedure of the MLP. The MLP was trained and validated using data collected in free space, while it was tested both in free space (introduced in Section 4.5.1) and in a constrained space (introduced in Section 4.5.2)

.

At each configuration, the tip position P_{Tip}, the tangent vector of the tip EM sensor $\overrightarrow{t_{Tip}}$, the base position P_{Base} and the base tangent vector $\overrightarrow{t_{Base}}$ can then be used to compute the ground truth curvature and the angle of the bending plane of all points where the gratings are presented along the catheter's length. Since the constant curvature model is valid, using the tip and the base pose together with the length l of the bendable segment, as shown next, it is possible to calculate the radius r and the center O of the circle that fits the catheter's shape. The distance between the tip and the base position is equal to the chord length c. The angle between $\overrightarrow{OP}_{Tip}$ and $\overrightarrow{OP}_{Base}$ is defined as α. The relations of these parameters are depicted in Fig. 4.4(a) and given by:

$$\alpha = \pi - \gamma, \tag{4.1}$$

$$r = \frac{c/2}{sin(\frac{\alpha}{2})}, \tag{4.2}$$

where γ is the angle between $\overrightarrow{t_{Tip}}$ and $\overrightarrow{t_{Base}}$. Finally, the curvature can be obtained by

$$\kappa = \frac{1}{r}. \tag{4.3}$$

Different pressure values were first applied to two antagonistic PAMs to make the catheter bend in two opposite directions. The method by Torr *et al.* [228] was used to fit a plane to all tip points. This plane is defined here as the so-called *zero-angle bending plane*. The normal vector of a plane that fits the catheter shape (which is named the current bending plane) is defined by taking the cross product of $\overrightarrow{t_{Base}}$ and $\overrightarrow{t_{Tip}}$. At each configuration, the angle of the bending plane θ is determined as the angle between the normal vectors of the zero-angle and the current bending plane. An example of the bending plane angle calculation can be seen in Fig. 4.4(b). The blue lines show the catheter shape during the process of deriving the zero-angle

Table 4.1: Mathematical functions of spiral (Sp) used during training and validation phase. Mathematical functions of Lissajous (Li) or circular (Ci) trajectories (Traj.) in the testing phase. All units are in millimeter.

Phase	Traj.	Mathematical functions	Traj.	Mathematical functions
Training& Validation	Sp	$x = 0.002\pi t\cos(0.04\pi t)$ $y = 0.002\pi t\sin(0.04\pi t)$		
Testing	Li1	$x = 0.6\cos(0.04\pi t)$ $y = 0.5\sin(0.08\pi t)$	Li2	$x = 0.8\cos(0.04\pi t)$ $y = 0.7\sin(0.08\pi t)$
	Ci1	$x = 0.4\cos(0.04\pi t)$ $y = 0.4\sin(0.04\pi t)$	Ci2	$x = 0.6\cos(0.04\pi t)$ $y = 0.6\sin(0.04\pi t)$

bending plane, while the black line shows an arbitrary current configuration of the catheter.

4.3.3 Data collection

To create the training data, two antagonistic PAMs were first sequentially pressurized to generate the zero-angle bending plane as described in Section 4.3.2. The catheter was then controlled to follow a 3D spiral trajectory (described in Table 4.1) that broadly cover the work space of the steerable catheter and that starts from the straight configuration. The training and validation data comprises 5880 samples from zero-angle bending plane and 22858 samples from the spiral trajectory, leading to a total of 28738 data points. The maximum curvature that was recorded during training was 19.3 m^{-1}. The 70% of collected data sets were used to train the ANN, while the rest 30% were used to validate the ANN. This was a random split. It was conducted using the “train_test_split” function in the “scikit-learn”, an open source machine learning library. The information on the training data is summarized in Fig. 4.5.

The proposed approach was tested with EM ground truth and image ground truth. Free space test sets with EM ground truth were created to verify the proposed approach by controlling the catheter’s distal tip to follow two circular and two lissajous trajectories. Note that the test trajectories are independent data sets that have completely different patterns with respect to the training set. By doing so, leakage from the training to the test set is avoided, ensuring proper validation of the approach. The mathematical functions of the trajectories are summarized in Table 4.1. Since the catheter was controlled in free space, the constant curvature assumption together with tip and base pose was used to generate the ground truth shape. However, as mentioned above, the constant curvature assumption is not valid in a constrained environment. To test the performance of the ANN in a constrained space, the ground truth shape of the catheter was generated from by camera images rather than using the catheter tip and base poses. The catheter was bent in a plane parallel to the camera in both a free and a constrained environment. During constrained environment tests, obstacles were placed in different places

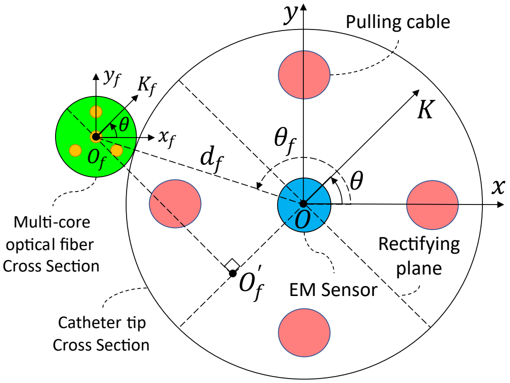


Figure 4.6: Cross sectional view of the catheter tip and the multi-core FBG fiber where θ is the angle of the bending plane with respect to the x_c-axis, d_f is the distance between the center of the catheter and the fiber's central axis, and θ_f is the angle between x-axis and the center of the fiber O_f.

along the catheter length. These obstacles caused the catheter to deform into non-constant curvature shapes.

4.3.4 Model-based off-center fiber calibration

Using a conventional model-based approach, a set of curvature κ_f and angle of the bending plane θ at discrete points along the fiber where the FBGs are inscribed can be calculated based on the geometry of the fiber core [229]. However, since the multi-core fiber is not placed on the central axis of the catheter, the calculated curvature does not indicate the catheter's curvature. The relation of the calculated curvature κ_f and the catheter's curvature κ is given by:

$$\kappa = g(\kappa_f) = \left({\kappa_f}^{-1} - \frac{\overrightarrow{OO_f'} \cdot \overrightarrow{OK}}{||\overrightarrow{OK}||} \right)^{-1}, \tag{4.4}$$

where O_f' is the projection of O_f (center of the fiber) on the bending direction vector $\overrightarrow{OK}$. The relative position of the multi-core fiber O_f and the center-line of the catheter O is characterized by the distance to the center-line d_f and the angle θ_f between $\overrightarrow{OO_f}$ and the x-axis. A cross sectional view of the multicore-fiber and the catheter tip is shown in Fig. 4.6.

An optimization problem is formulated to find d_f and θ_f by minimizing the following objective cost function.

$$\arg\min_{d_f,\theta_f} \frac{1}{n \times m} \sum_{i=1}^{n}\sum_{j=1}^{m} \|g(\kappa_{f_{i,j}}) - \kappa_{gt_i}\|, \tag{4.5}$$

where m is the number of points where FBGs are inscribed along the length of the optical fiber, n is the number of samples in the training data, and κ_{gt_i} is the ground truth curvature that is generated from the EM poses. Two algorithms supported by MATLAB, namely *trust-region reflective* [230] and *Levenberg-Marquardt* [231] were used to solve the optimization problem. The result of *trust-region reflective* was chosen since it yields a lower value for the objective function (approximately 15%).

4.3.5 Shape reconstruction

As explained in the following, the discrete estimated curvature of the catheter central axis and the angle of the bending plane are interpolated. This ensures the continuity of the estimated curvature and improves the accuracy of shape reconstruction. The catheter shape is represented as a space curve, which can be described as a set of curvatures $\kappa(s)$ and torsion $\tau(s)$, where s is the variable arc length. Torsion $\tau(s)$ is the rate of change of angle of the bending plane with respect to the arc length. The curvature and torsion profiles determine the evolution of the tangent $\overrightarrow{t}$, the normal $\overrightarrow{n}$, and the binormal $\overrightarrow{b}$ unit vectors along the length of the catheter. By solving the differential Frenet-Serret formulae, the evolution of the TNB frame is obtained [232]. The Cartesian position of each point that belongs to the catheter length $C(s)$ follows then by integrating

$$C(s) = \int_0^s \overrightarrow{t}\, ds. \tag{4.6}$$

In this work, (4.6) is solved using the Helical Extension Method (HEM) [233] as suggested in [217].

4.4 ANN-based off-center fiber shape reconstruction

4.4.1 Design of the ANN

In this work, a multilayer perceptron (MLP) is used to auto-calibrate the shape sensing functionality of the robot. A multilayer perceptron is a class of feed forward artificial neural network that consists of an input layer, several hidden layers and an

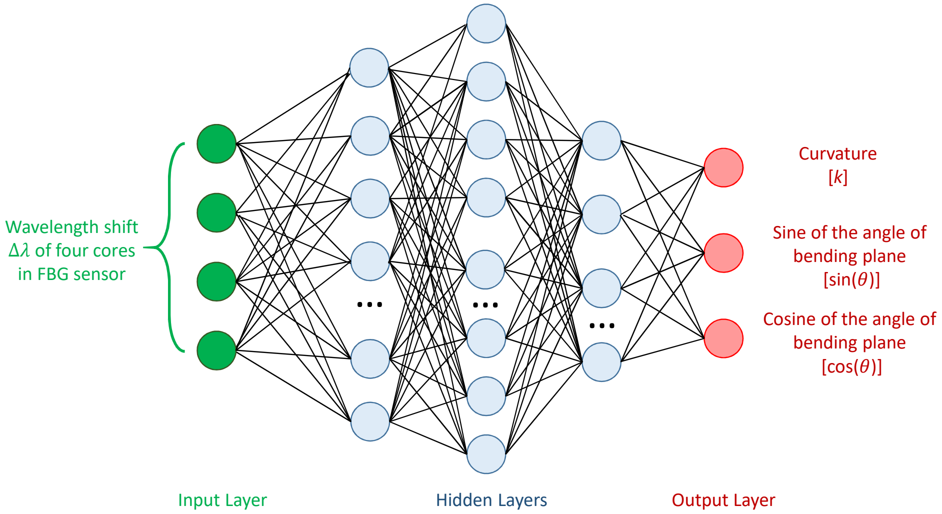


Figure 4.7: A multilayer perceptron (MLP) for FBG calibration. The four wavelength shifts $\Delta\lambda$ from the four-core FBG fibers are entered into the input layer of the proposed MLP. The output layer has three neurons. One neuron predicts the curvature, while the other two predict the sine and the cosine of the angle of the bending plane. Five of these MLPs are established for the five gratings that are distributed over the entire length of the catheter segment.

output layer [234]. Except for the neurons in the input layer, each neuron adopts a nonlinear activation function. These nonlinear activation functions turn the MLP into a nonlinear perception, thus distinguishing itself from linear approaches, such as linear regression. Thanks to a large number of neurons and the possibility to make use of multiple types of nonlinear activation functions, MLPs can model complex nonlinear behavior [235]. MLP training is based on a supervised learning algorithm called back-propagation [236]. The training procedure requires sufficient data containing various patterns. The MLPs developed here relate the measured wavelength shifts from FBGs directly to the curvature and angle of the catheter's bending plane without explicit knowledge of the correspondence between the fiber's position and the catheter center-line or explicit knowledge of the characteristic parameters of the fiber. Note that the reason for using wavelength shifts $\Delta\lambda$, instead of raw wavelengths values, is that the absolute value of the wavelength itself could be influenced by environmental factors such as temperature and humidity, which is less for the case of wavelength shifts.

The length of the flexible part of the catheter segment is 40 mm. Since the distance between neighbouring gratings is 10 mm, only five sets of curvature and angle of the bending plane is needed to reconstruct the shape. Therefore, five MLPs, as shown

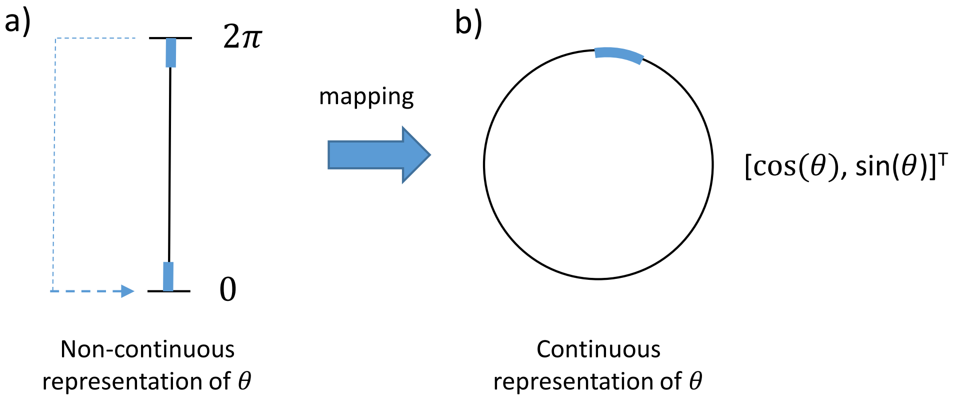


Figure 4.8: Instead of directly estimating the angle of the bending plane, the proposed MLP divides it into two outputs, i.e. the sine and cosine of the angle of the bending plane. This approach transforms the discontinuous rotation representation *(left)* into a continuous representation *(right)*, thus facilitating the fitting of the neural network.

in Fig. 4.7, were employed. Here, rather than a large MLP that takes five gratings together, five separate MLPs were trained. This allows efficient training and helps transferability of the results to more general scenarios. Training can be done in free space only (which is faster than training every possible contact situation), but - as demonstrated in this work - still allows learning the behavior/parameters of each separate grating. This is found helpful to generalize to more general cases where the shape deviates from a single constant curvature for all the five gratings (as would be expected in free space).

Each MLP automatically interprets the strains from the four gratings of an individual location. Each MLP (see Fig. 4.7) consists of five layers i.e. one input layer, three hidden layers and an output layer. The input layer has four neurons, which receive the wavelength shifts $\Delta\lambda$ that are collected from a four-core FBG fiber. The catheter was first placed in a straight configuration to measure the unstrained Bragg wavelength λ_0 of each grating. The wavelength shifts were then calculated as:

$$\Delta\lambda = \lambda - \lambda_0, \tag{4.7}$$

where λ is the strained Bragg wavelength.

The first, second and third hidden layer have 180, 360, and 90 neurons, respectively. A Rectified Linear Unit (ReLU) is used as the activation function and added after each hidden layer. The output layer has three neurons. One of the outputs is the curvature. The two other neurons predict the sine and the cosine of the angle of the bending plane. This allows circumventing the discontinuity of the angle at 0 and 2π (Fig. 4.8(a)) that is otherwise difficult to cope with for an ANN [237]. The prediction results of the MLP are shown in Fig. 4.9(a) for the case where

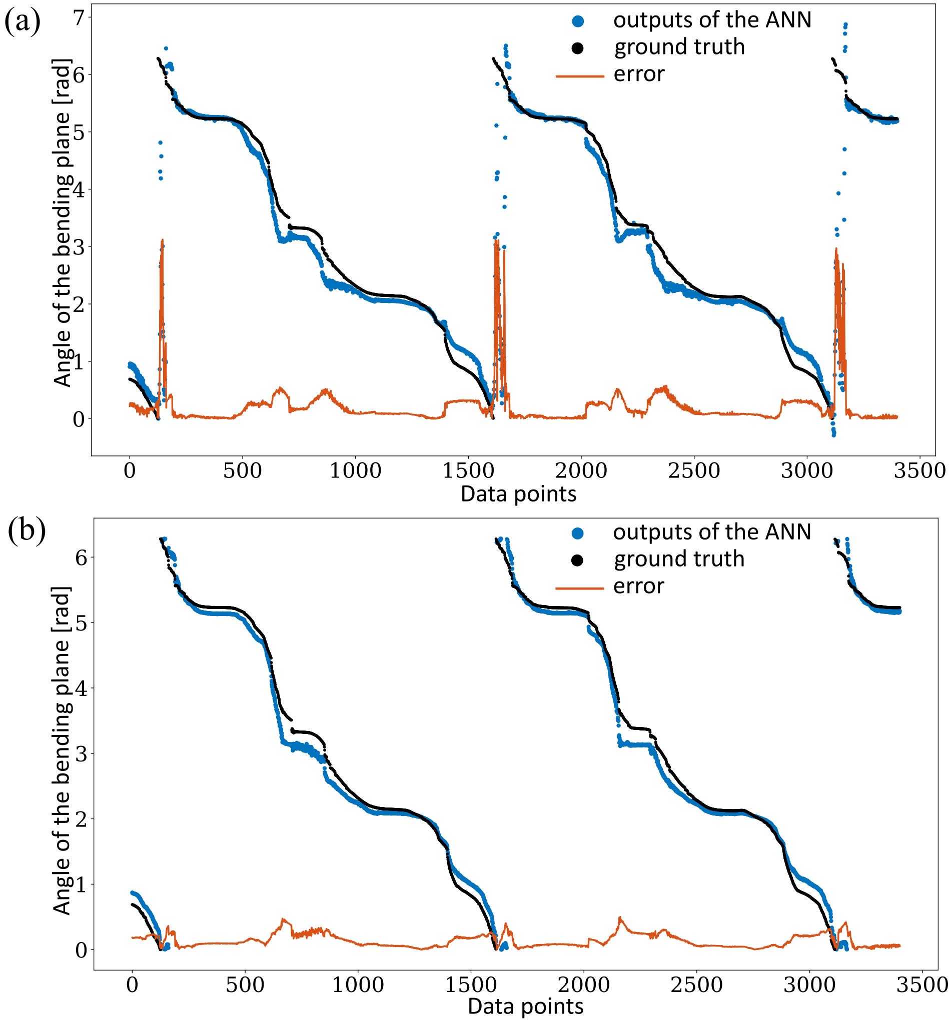


Figure 4.9: Prediction of the angle of the bending plane of grating 3 using different representation methods: (a) non-continuous representation: as can be seen in the dashed green boxes, the MLP with the non-continuous representation generates large errors in the transition region $[0{,}2\pi]$; (b) continuous representation: MLP performs well in the transition region, thus produces small errors.

Table 4.2: Feature Ablation Study

	Avg. validation loss (std)	Percentage increase
remove the 1st core	0.0065 (0.0004)	14.0%
remove the 2nd core	0.0081 (0.0003)	42.1%
remove the 3rd core	0.0079 (0.0006)	38.6%
remove the 4th core	0.0069 (0.0001)	21.1%

Table 4.3: Hyperparameters for the MLP network

	Number of hidden layers	Number of neurons per cell	Activation functions	Optimizer	Loss function
MLP	3	180, 360, 90	ReLu	Adam	L2 Loss
	Training-subset/Validation ratio	Batch size	Learning Rate	Epoches	
	70%/30%	16	10^{-5}	350	

the non-continuous representation has been used. As visible in the green dashed boxes, there are many haphazard points distributed over the transition region [0, 2π]. The error plot (in green) in this region generates a peak, where the error is significantly larger than the other regions. In this case, the Mean of Absolute Error (MoAE) of the 3^{rd} set of gratings is 0.21 rad. If we reconfigure and train the ANN with its output to the sine and cosine of the angle of the bending plane, the rotation is represented in a continuous fashion (Fig. 4.8(b)). This will improve the learning efficiency of the MLP (see Fig. 4.9(b)). The MoAE obtained in this case (results shown in Fig. 4.8(b)) is 0.13 rad, which is 38.1% better than the non-continuous representation. The angle of the bending plane can be finally calculated by wrapping the angle back to the interval of 0 to 2π via the operation:

$$\theta = \text{atan2}\left(\frac{\sin\theta}{\cos\theta}\right). \tag{4.8}$$

To increase the learning efficacy, all the input and output training data were normalized between [-1,1] following (4.9).

$$x_{norm} = (b-a)\frac{x - x_{min}}{x_{max} - x_{min}} + a, a = -1, b = 1. \tag{4.9}$$

The output normalization regulates the three outputs to the same range, thus the ANN learns the three outputs with equal efforts. Here, a normalization is carried out instead of a standardization because normalization is suitable for use when the distribution of data does not follow a Gaussian distribution. In addition, ANN is an algorithm that does not assume/prerequire any type of data distribution. Note that in (4.9), the variable x represents respectively the wavelength shift $\Delta\lambda$, curvature κ, and the sine and the cosine of the angle of the bending plane: $sin(\theta)$ and $cos(\theta)$, respectively.

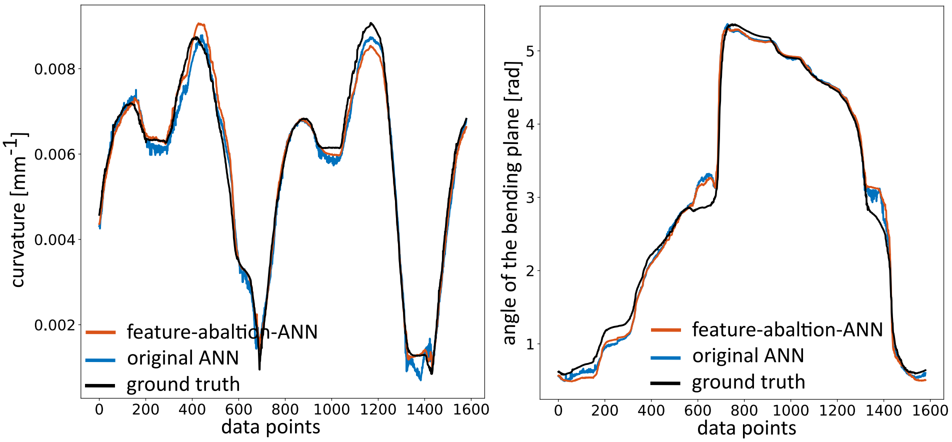


Figure 4.10: The results of the feature ablation study. The feature-ablation-ANN (with three inputs) and the original ANN (with four inputs) trained for grating set 3 were evaluated on testing trajectories Li1 in Table 4.1. The feature-ablation-ANN could achieve a comparable performance as the original ANN even if the information from the central core is not provided.

4.4.2 Feature ablation study of the ANN

A feature ablation study was conducted to determine the contribution of each input feature to the overall performance of the ANN. Ablations were carried out by removing one of the four inputs. This means that the ANN used for the feature ablation study (hereinafter referred to as “feature-ablation-ANN”) only has three inputs. Next, the feature-ablation-ANN was re-trained. To investigate the effect of the ablation, the performance of the feature-ablation-ANN was first evaluated on the validation set. The validation losses are reported in Table 4.2. It shows that removing the first core (central core) leads to the lowest increase (14.0%) in validation loss. Because, the central core practically coincides with the fiber’s neutral axis, it is not sensitive to bending strain. Therefore, it provides less information to the ANN compared to other cores and removing it has not that large impact.

The feature-ablation-ANN (with central core removed) was further validated on the test set. An ANN was trained for each grating. Five feature-ablation-ANNs are applied to the four testing trajectories. The errors of curvature and the angle of the bending plane generated by the feature-ablation-ANN were compared to that of the original ANN. One group of results tested on the trajectory Lissajous (1) is shown in Fig. 4.10. The results show that the average curvature error of the feature-ablation-ANN (0.56 m^{-1}) is slightly smaller than the original ANN (0.62 m^{-1}). Their average errors in terms of the angle of the bending angle are comparable, namely 0.20 rads for feature-ablation-ANN and 0.21 rads for original

ANN. This can be explained by the fact that the central core provides redundant information. The ANN could achieve better performance when the redundant input information is removed. However, it is expected that the information from the central core could be beneficial in more complex loading cases such as a longitudinal load, and/or a variable temperature. These disturbances have not been tested at this point and remain the subject of further work. Based on this consideration, this work continues with the original ANN. Note that the aim of the feature ablation study described here is to find efficient combinations of hyperparameters and ANN structures. The aim is not to find the optimal configuration for a certain data-set, as this would not necessarily guarantee optimal performance on other datasets.

4.4.3 Summary of hyperparameter study

The hyperparameters of each MLP are listed in Table 4.3. Hyperparameters, e.g. Adam optimizer [44], training/validation ratio 30%/70% [238, 239], or ReLU activation function [44], were selected based on previous research or empirically. The network was implemented using Python, and in particular PyTorch, an open source machine learning library. Ubuntu 20.04 was the operating system that was used. Each MLP has been trained for 350 epochs until the training and validation loss reached a plateau. The whole training time was less than 15 minutes. The loss trend of one training session was shown in Fig. 4.11. The training results indicate that the MLP converged at the end of the training process. The average inference time of the MLP for a single output is around 0.18 ms. Also considering the MLP has a practical structure among the ANNs, it is feasible for the users to train an MLP in a short time. The users can employ the trained MLP to achieve the calibration of the FBG as well as the reconstruction of the robot shape in real-time.

4.5 Results and discussion of the DL-based improved catheter shape sensing

Figure 4.5 summarizes the verification procedure of the trained MLP. The MLP was first verified in free space by reconstructing the shape of the robot, while it was controlled to follow two Lissajous and two circular trajectories. Subsequently, verification took place in a constrained space. Here, the catheter was brought into contact with an obstacle at different locations along its length.

4.5.1 Free space experiments with ground truth generated EM

In the free space experiment, with ground truth generated by EM, the catheter was commanded (in open loop) to roughly follow the predefined trajectories as

Table 4.4: Experimental results for ANN and model-based approach in free space, ground truth generated by EM (first four rows), and constrained environment, ground truth from camera (last seven rows). The individual error for each constrained experiment Fig.4.14(b)-(g) are also reported.

Experiment		Mean error for one shape at a single time step (mm)		Maximum error for one shape at a single time step (mm)		Overall maximum error (across all shapes) (mm)	
		ANN	Model	ANN	Model	ANN	Model
EM ground truth	lissajous (1)	0.158 ± 0.090	0.250 ± 0.130	0.370 ± 0.213	0.746 ± 0.372	1.080	1.562
	lissajous (2)	0.216 ± 0.088	0.304 ± 0.160	0.542 ± 0.205	0.882 ± 0.503	1.009	1.644
	circular (1)	0.285 ± 0.086	0.352 ± 0.259	0.782 ± 0.282	1.038 ± 0.789	1.773	2.148
	circular (2)	0.309 ± 0.165	0.335 ± 0.156	0.838 ± 0.380	0.945 ± 0.497	1.795	3.007
Camera ground truth	(a)	0.491 ± 0.053	0.569 ± 0.151	0.624 ± 0.096	0.721 ± 0.203	0.851	1.170
	(b)	0.507 ± 0.048	0.621 ± 0.039	0.711 ± 0.055	0.572 ± 0.171	0.839	0.926
	(c)	0.460 ± 0.028	0.641 ± 0.039	0.708 ± 0.073	0.859 ± 0.071	0.816	0.938
	(d)	0.503 ± 0.021	0.565 ± 0.029	0.665 ± 0.038	0.761 ± 0.052	0.756	0.819
	(e)	0.483 ± 0.018	0.405 ± 0.133	0.632 ± 0.027	0.533 ± 0.161	0.658	0.723
	(f)	0.499 ± 0.016	0.510 ± 0.016	0.603 ± 0.040	0.627 ± 0.027	0.675	0.678
	(g)	0.511 ± 0.021	0.654 ± 0.024	0.701 ± 0.037	0.778 ± 0.039	0.800	0.848

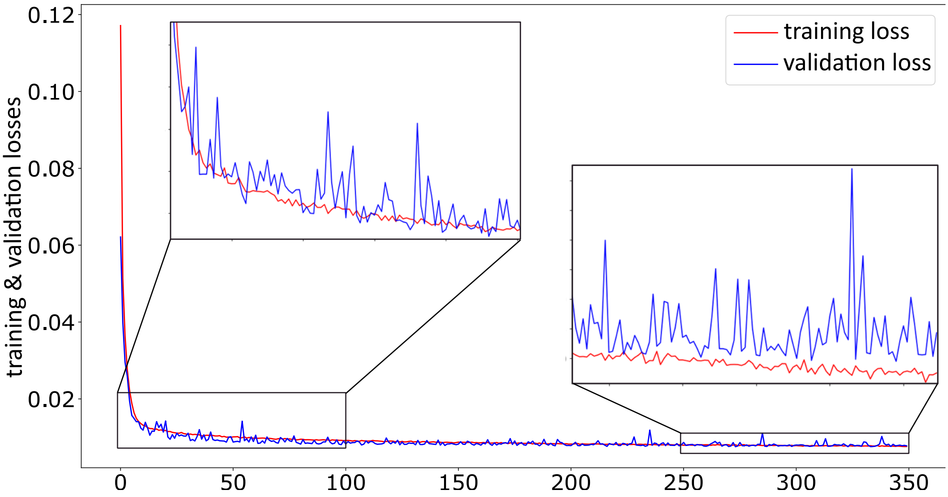


Figure 4.11: Loss trend (grating 3) presented as training loss (red) and validation loss (blue) vs. the number of epochs.

shown in Table 4.1. The catheter bends with curvatures up to 15 m^{-1}. Note that no effort was made to follow the said trajectories closely (*i.e.*, open loop control was implemented) as this work focuses on reliably predicting catheter shape, not on closely following a trajectory. The ground truth of these experiments was generated using the tip and base poses together with the constant curvature assumption of the catheter. The error as a function of the arc length is defined as the Euclidean distance between the reconstructed and the ground truth data:

$$E_{dist} = d_{min}(S_{gt}, S_{est}), \tag{4.10}$$

where S_{est} is the reconstructed shape expressed as a set of k points $S_{est} = [s_1 s_2 ... s_k]$ and $s_{i \in \{1,...,k\}}$ is the 3D coordinate of each point. The distance error E_{dist} is the distance between the i^{th} point of the estimated shape S_{est} to the closest point on the ground truth shape S_{gt}.

The blue lines, which are shown in the graphs in the first row of Fig. 4.12(a-d), depict the estimated catheter shapes using the ANN approach. Distance errors are shown in color-coded map graphs in the second row of Fig. 4.12 in panels (e-h). The histograms showing the mean and maximum error are shown in Fig. 4.12(i-f). The mean error for one shape is calculated by averaging E_{dist}, while the maximum error shows the highest value in E_{dist} at each time step. The experimental results for the free space with ground truth EM are shown in the first part of Table 4.4. In this table, the mean and maximum error for each shape at a single time step using our proposed ANN approach and the traditional model-based approach are reported with their standard deviation. The overall maximum error that shows the largest error in each experiment is also mentioned in Table 4.4. One can

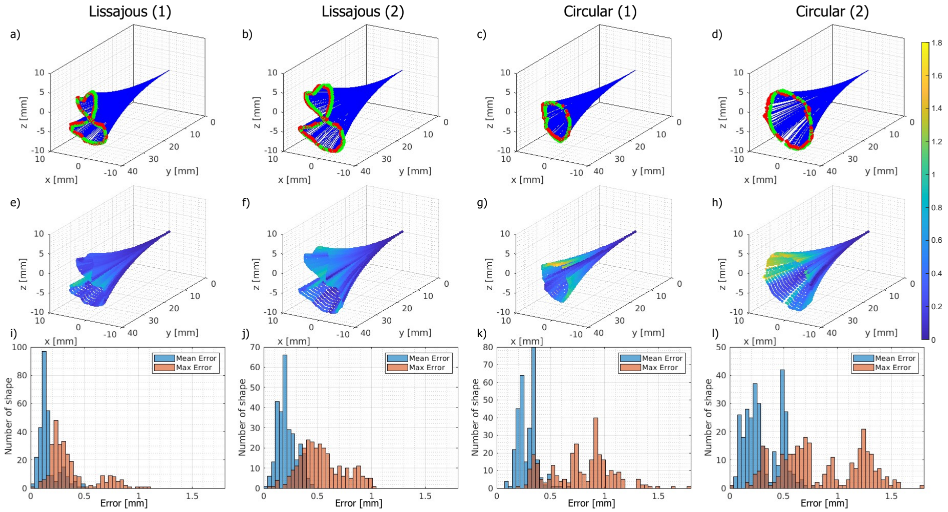


Figure 4.12: The results of free space shape sensing experiments with EM ground truth using ANN approach: (a)(e)(i) Lissajous (1); (b)(f)(j) Lissajous (2); (c)(g)(k) circular (1); (d)(h)(i) circular (2). The catheter reconstructed shapes are presented by blue lines in the graphs displayed on the first row while the images on the second row show the color-coded Euclidean distance errors of each point along the catheter length. The images on the third row show the histogram of the mean and maximum error for each trajectory.

observe that the proposed approach can estimate the catheter shape with good accuracy with an overall mean and maximum reconstruction error of 0.244 mm and 0.639 mm, respectively. The color-coded map graphs (Fig. 4.12(e-h)) show that the maximum error typically appears at the tip of the catheter. This is logical, as errors are integrated along the catheter length, and therefore the maximum error can be considered as the estimated tip error. The catheter experienced different bendings in the four different trajectories. The largest mean error of 0.309 mm appears at the circular trajectory (2) where also the largest curvature takes place. The overall maximum errors of the two Lissajous trajectories were 1.080 mm and 1.009 mm, while the maximum errors of the two circular trajectories were 1.773 mm and 1.795 mm.

Figure 4.13 shows the mean error over the time of Lissajous (1) experiment together with the estimated angles of the bending plane for each grating and the ground truth angles of the bending plane. The mean error increases at the transition points of the angle bending plane while the catheter crosses the zero-angle of the bending plane (marked with dashed squares in Fig. 4.13). Note that here the zero-angle of the bending plane is calibrated to be aligned with the switching point of the two antagonistic muscles (Fig. 4.6) of the catheter. The large error at the

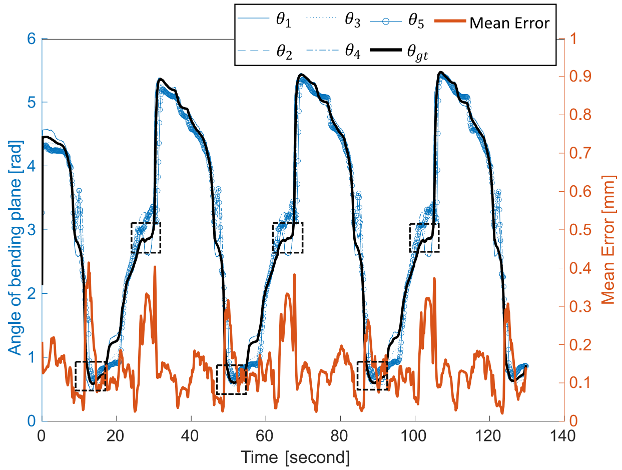


Figure 4.13: The estimated angles of the bending plane of each grating (blue), ground truth angle of the bending plane (black) and mean error (orange) of lissajous (1) experiment are plotted over the time. Transition regions where there are increases in mean error are marked with dashed square.

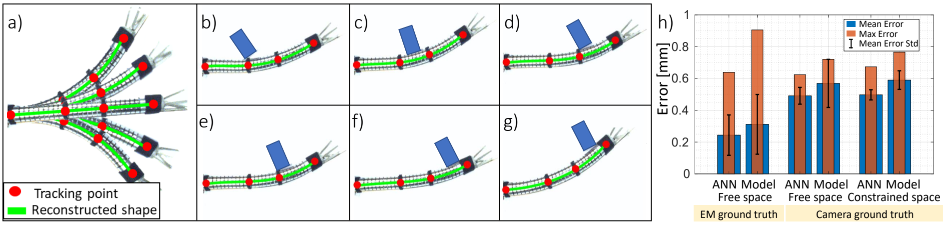


Figure 4.14: The results of constrained environment shape sensing experiments using the ANN approach: (a) free space; (b)-(g) constrained space. The catheter reconstructed shapes are presented by a green line while the tracked markers are shown in red. The quantitative shape reconstruction results of the experiments with EM ground truth and with camera ground truth using ANN and model-dependent approach are presented in (h).

transition point is due to the fact that the low frequency of the EM sensing is not able to capture fast dynamic movement of the catheter tip when it crosses the switching points of the two antagonistic muscles. The unavoidable non-linearities (e.g. hysteresis, dead zone) make the catheter tip movement more unpredictable. This can be clearly seen in Fig. 4.9. Consequently, the quality of the training data collected at the transition point is worse. In turn, this leads to lower accuracy, here seen as a larger error of the bending plane angle around the zero-angle bending plane.

For comparison, the catheter shape was also reconstructed with a conventional model-based shape sensing method introduced in [229] with the calibration step described in Section 4.3.4. The mean and maximum reconstruction errors achieved by this conventional method are 0.311 mm and 0.906 mm, respectively, which are 21.5% and 29.5% larger than the results achieved by the ANN-based method (Fig. 4.14(h)).

4.5.2 Free space and constrained space experiments with ground truth generated by the camera

In the experiments with ground truth generated by the camera, two antagonistic PAMs were successively actuated to make the catheter bend in a plane parallel to the camera's plane. This allows us to use the camera as a source for ground truth generation. To generate the constrained environment test sets, rigid obstacles were fixed at various locations along the catheter length as shown in Fig. 4.14. During these experiments, the PAMs were pressurized with a maximum value of 3.5 bar which results in the maximum recorded curvature of 20 m^{-1}. Given the short length of the bendable segment (approximately 40 mm) and the large size of vessels where we are targeting at, the segment would only assume to get into S-shape configuration when an intense contact appears at the tip. Such contact would need to be maximally avoided. For these reasons, we did not include S-shaped bending in evaluating our proposed approach.

By controlling the catheter to move in the image plane, the camera was able to capture the entire shape of the catheter. The images recorded by the camera were used to generate ground truth for the unconstrained and the constrained environment experiments since the presence of extrinsic disturbances (obstacles) makes the catheter deform into shapes for which the constant curvature assumption is no longer valid. The interaction with an obstruction complicates the catheter shape segmentation process. To address this issue, the heat shrink's positions along the catheter length were tracked rather than the centerline of the entire catheter. Color segmentation was first applied to the input image to identify the contours of the heat shrink. The heat shrinks' locations in the image frame $P^{Img}_{m\times 2}$ (where m is the number of heat shrink) were then determined by the center of the recognized contours. Since the center points of the heat shrinks were moving in a plane during the experiments, the 2D positions of these points can be converted to 3D positions by adding a fixed z-value (i.e. $z = 0$). The 3D positions of the heat shrinks in the ground truth frame were defined as $P^{GT}_{m\times 3} = scale \times [P^{Img}_{m\times 2}\, 0_{m\times 1}]$ where *scale* is a camera scale factor calibrated by observing a known-size chess board. To register the FBG shape reconstruction frame and the ground truth frame, a spatial calibration step as detailed in [229], was performed to find the correspondence between the measured shape and the measured positions of the four heat shrinks. The catheter was fixed at the base and bent in two configurations symmetric with respect to the straight configuration. The distance between the first

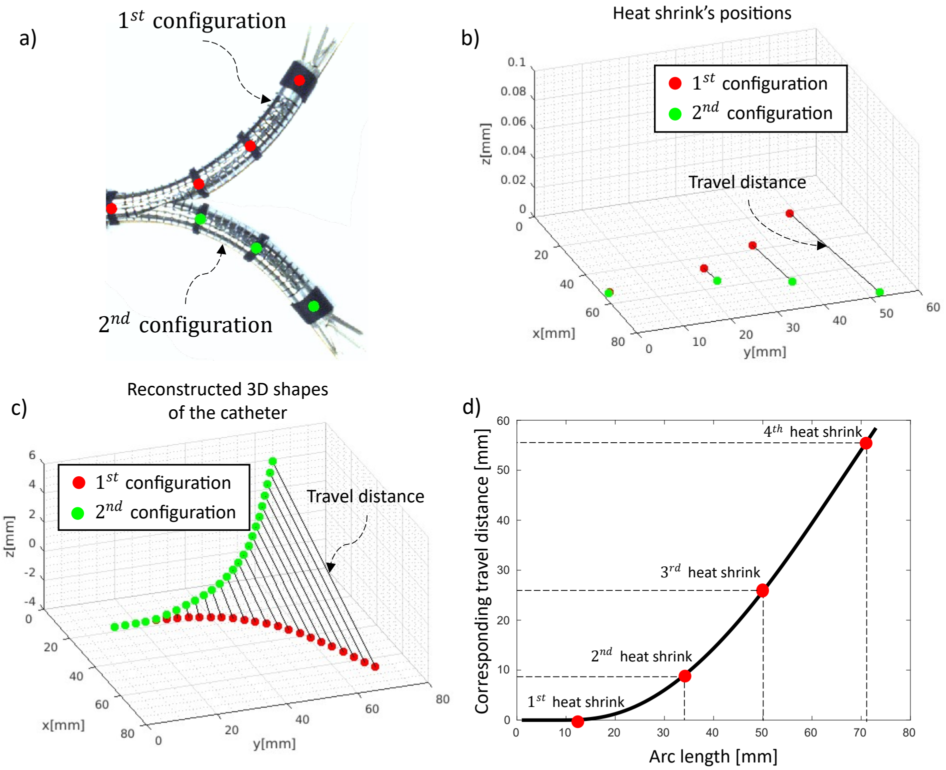


Figure 4.15: The spatial calibration step to find the correspondence between the heat shrinks position and the 3D reconstructed shape. Catheter was first bent in two configurations as shown in (a). The heat shrink's positions were then extracted by image processing and transform to the ground truth coordinate frame as can be seen in (b). The 3D shapes of the catheter in two configurations were reconstructed in the FBG coordinate frame (c). The travel distance between two configurations of each point along the catheter length was calculated and compared to the travel distance of each heat shrink to find the correspondence (d).

and second configurations of each corresponding point at a given arc length along the reconstructed shape was measured. The travel distance of each point along the length of the reconstructed shape was then compared to the travel distance of each heat shrink to find the corresponding arc length of each heat shrink. An example of this spatial calibration process is shown in Fig. 4.15.

To localize the 3D reconstructed shape in the ground truth frame, m shape points in the FBG shape reconstruction frame together with their corresponding points in the ground truth frame ($P^{GT}_{m\times3}$) are used to find a rigid transformation matrix at each time step. In this work, a point-to-point registration method proposed by Horn [240] was used. To compare the efficiency of the ANN-based and the model-based approach, the mean and the maximum distance between the four heat shrink points and their corresponding points in the estimated shape in the ground truth coordinate frame were calculated.

Figure 4.14(a) shows the estimated shape in the unconstrained experiments while Fig. 4.14(b-g) visualizes the catheter shape in the constrained environments. The experimental results of the free space and of the different contact location experiments, with camera ground truth, are summarized in the second and third parts of Table 4.4. The quantitative shape reconstruction results using the ANN and the model-based approach are shown in Fig. 4.14(h). The mean, maximum value and the standard deviation of the shape reconstruction error are plotted in the constrained environment with EM ground truth, free environment and constrained environment with camera ground truth. It can be observed that the proposed ANN approach yields a mean shape deviation of 0.491 mm and 0.497 mm in the free and constrained environment experiments with camera ground truth, respectively (Fig. 4.14(h)). The traditional model-based approach shows a similar consistent accuracy of approximately (0.568 mm) in both free space and constrained space experiments. In terms of the maximum reconstruction error, the ANN improved the performance over the traditional approach by 13.9% and 11.8% in free space and constrained space, respectively.

4.5.3 Discussion on the results of free space and constraint space experiments

Referring to Fig. 4.14(h), one can note that the maximum error in the free space experiment with EM ground truth is larger than the maximum error of the experiments with camera ground truth. This is due to the fact that the evaluation metrics in the EM ground truth experiments (pointwise shape reconstruction error over the entire length) are different from those used in the camera ground truth experiment (shape reconstruction error using four heat shrinks as reference points). Therefore, the errors between these two experiments represent two different metrics and are not comparable. These experimental results reflect the superiority of the proposed ANN approach compared to the traditional model-based approach in all three experiments. It is worth mentioning that due to the fast response and the high frequency streaming rate of the FBGs, the shape reconstruction algorithm runs at 200 Hz using MATLAB on a computer with specifications mentioned in Section 4.3.3.

Considering the average error and the maximum shape sensing error, the authors believe that this method outperforms the results reported in [229]. In addition, our method only requires data acquisition in free space to train the ANN, and the trained ANN can be generalized to the constrained space as well. The ANN efficiently learns the mapping from the wavelength shift $\Delta\lambda$ to the curvature and the angle of the bending plane. This mapping is in fact an intrinsic property of the robot. Indeed, it is shown through these experiments that this mapping does not depend on whether the robot is in contact with obstacles or not. Therefore, it can be said that the ANN model trained in free space is well transferable to contact situations. Aside from the elegance and knowledge that free space training suffices, the training itself is, with 30 minutes training in total, also very fast. One would expect this training time to rise proportionally for longer sections, but training would only be required once, e.g. right after construction of the catheter.

In addition, considering the concerns arising from the black-box nature of DL methods, two precautions are proposed in case the ANN fails: 1) the maximum curvature that can be achieved by the flexible robot is identified before interventions. If the prediction of the ANN is greater than this threshold, or less than zero, then the system should indicate to clinicians that the reconstructed shape is unreliable; 2) the traditional model-based approach that is used as a comparison to the ANN will be used in parallel with the ANN. If the situation described in 1) occurs, or the difference between the model-based approach and the ANN is too large, or if the ANN has no outputs at all, then the system switches to the traditional model-based approach for shape reconstruction. This mechanism takes advantage of the superior accuracy of the ANN, but at the same time offers redundancy to mitigate the risks when the ANN approach would fail during interventions.

4.6 Conclusion on DL-based improved catheter shape sensing

In this chapter, a new data-driven approach to auto-calibrate the shape sensing functionality of a multi-DOF catheter is presented. The newly presented solution allows the optical fiber to be inserted in an off-center channel while still ensuring a high shape sensing precision. To prove the superiority of the proposed data-driven approach compared to the traditional model-based approach, both methods were implemented in this work. A method to map the curvature measured by the fiber inserted to an off-center channel to the catheter's curvature is also presented. This allows a fair comparison between the two approaches. The performance of the proposed method was evaluated with different ground truths in both free and constrained environments and yielded promising shape sensing accuracy. Using the proposed method, average shape sensing errors of 0.24 mm and 0.49 mm were recorded in free space and constrained space experiments, respectively.

Currently, the training data sets were generated by the catheter tip and base poses with constant curvature assumption. In the future, other sensing modalities such as multi-view 3D computer vision method named shape reconstruction from silhouette [241] will be considered for use in training data generation. The use of shape reconstruction methods from silhouette will enable our method to be also applied to different types of flexible instruments that do not necessarily need to follow the constant curvature assumption in free-space during training as was done in this chapter.

This chapter addressed the issue of limited visual feedback in Chapter 2, advancing to showing the 3D shape of the catheter instead of presenting solely the catheter tip's pose. DL methods were employed to enhance shape sensing in continuum robots using a non-radiative approach. During data collection and validation, the hysteresis effect is evident. As observed in Fig. 4.10, there is a lack of response in the curvature and the angle of the bending plane when pressure is varied. Such phenomena could influence the precision of catheter tip positioning, which could potentially lead to risks of tissue damage or even perforation. No matter using teleoperation as shown in Chapter 3 or (semi-)autonomous control as showcased in Chapter 6, the hysteresis needs to be understood and compensated. The following chapter introduces a DL method to model hysteresis in robotic catheters. Its ability to enhance reconstruction of anatomical environments is also showcased when combined with forward-looking imaging modalities.

Chapter 5

Hysteresis modeling of robotic catheters based on a Long Short-Term Memory network for improved environment reconstruction

This chapter is based on the following publications:

D. Wu*, Y. Zhang*, M. Ourak, K. Niu, J. Dankelman, E. Vander Poorten, "Hysteresis Modeling of Robotic Catheters based on Long Short-Term Memory Network for Improved Environment Reconstructio," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2106-2113, April 2021

D. Wu, A. Sridhar, W. Beckers, S. Z. Mehdi, M. Ourak, I. Tamadon, J. Dankelman, E. Votta, A. Menciassi, E. Vander Poorten, "Data-driven Modeling of Complex Hysteresis Behaviour in MitraClip Steerable Catheters," *12th Conference on New Technologies for Computer and Robot Assisted Surgery*, September 2023.

The positioning precision of the catheter tip is negatively affected by hysteresis. To ensure that tissue damage is avoided due to imprecise positioning, hysteresis must be understood and compensated for. This work investigates the feasibility to model hysteresis with a Long Short-Term Memory (LSTM) network. A bench-top experimental setup containing a catheter distal segment was developed for model evaluation. The LSTM was first tested using four groups of test datasets containing diverse patterns. To compare with the LSTM, a Deadband Rate-Dependent Prandtl-Ishlinskii (DRDPI) model and a Support Vector Regression (SVR) model were established. The results demonstrated that the LSTM is capable of predicting the tip bending angle with sub-degree precision. The LSTM outperformed the DRDPI model and the SVR model by 60.1% and 36.0%, respectively, in arbitrarily varying signals. Next, the LSTM was further validated in a 3D reconstruction experiment using Forward-Looking Optical Coherence Tomography (FL-OCT). The results revealed that the LSTM was able to accurately reconstruct the environment with a reconstruction error below 0.25 mm. In addition to this, we have also verified the capability of the LSTM to handle hysteresis patterns in another system that features different hysteresis patterns and expanded range of motion. Overall, the proposed LSTM enabled precise free-space control of a robotic catheter in the presence of severe hysteresis. The LSTM predicted the catheter tip response precisely based on proximal input commands, minimizing the need to install sensors at the catheter tip for localization.

In Chapter 3, the teleoperation of the robotic catheter is realized using an AR-HMD and a gamepad. Chapter 4 presents the use of a custom-made, pneumatically-actuated catheter for collecting the requisite training and test data, which is then used to train an ANN aimed at improving the accuracy of FBG-based shape reconstruction. Despite these advances, the hysteresis effect remains uncompensated in above scenarios. The hysteresis effect impedes intuitive catheter control and consequently reduces the control precision of the catheter. This chapter addresses Sub-objective 4 introduced in Section 1.7 and delves into the critical issue of hysteresis in catheters. Addressing hysteresis is crucial to enhance the intuitiveness of physicians during teleoperations. It also plays a crucial role in increasing the positioning accuracy of the catheter tip, which is needed in the long term if one wants to automate the procedures.

5.1 Introduction

Robotic catheters can be operated based on various working principles [242]. Cable-driven technology is one of the most popular driving principles for robotic catheters. The cables, which are routed over the entire length of the catheter, undergo quite some friction with their guiding tubes. Consequently, reaching a high bandwidth is generally difficult. This hinders the deployment of cable-driven technology in scenarios in which good responsiveness is desirable. From the 1950s onward, Pneumatic Artificial Muscle (PAM) became popular for many applications due to its high operation bandwidth, easy fabrication, and low-cost [243]. PAMs also show good promise for being used in interventional instruments, in which precise maneuvers are required [244].

5.1.1 Prior works

Whether actuated by cables or by PAMs, precise control of catheters or, in a broad sense, flexible surgical robots is challenging. With hysteresis being a primary source of imprecision, hysteresis establishes a complex non-linear multi-valued relation between input commands and the response of the catheter tip. This multi-valued relation complicates navigation and results in inaccurate positioning of the catheter tip. To solve this problem, several methods have been explored in the past (Table 5.1). The use of external sensors as feedback to implement a closed-loop control was presented in [245], [246]. However, mounting sensors at the distal end is challenging due to spatial restrictions and sterilization requirements [247]. Some previous research also demonstrated the feasibility of using imaging techniques as feedback to minimize hysteresis [248], [249]. Baek *et al.* further integrated these imaging-based methods with kinematic models [133]. Amongst others, analytic models for identifying hysteresis is one of the most popular research fields ([250, 251, 252]). An open-loop controller could be designed based on the inversion of the identified models. Nevertheless, analytic models require a large number of parameters, which would lead to a cumbersome identification process. It is worth noting that, except for [252], all the above-mentioned research tackled hysteresis in cable-driven systems, while the study of hysteresis in fluidics-driven systems was scarce.

Deep learning techniques gained interest as they allow to avoid intricate identification processes by training an artificial neural network. Up to now, few works have attempted to use machine learning or deep learning for hysteresis modeling. Xu *et al.* [253] employed regression methods to learn the inverse kinematics model of a serpentine surgical manipulator. Porto *et al.* [247] used machine learning to produce position control of a flexible surgical robot. Both works dealt with cable-driven robots, and adopted traditional machine learning methods but did not take advantage of recent advances in deep learning. Several researchers used various deep learning approaches to model force hysteresis targeting generic [254], [255] and medical applications [221], [89]. In this work, we tackle hysteresis in positioning tasks, whereas force tasks - as described in [89, 221, 254, 255] form the topic of furher investigations and are out of the scope of this work.

5.1.2 Chapter contributions & structure

The chapter proposes the use of a Long Short-Term Memory (LSTM) network to model hysteresis in a catheter system. It is shown that the proposed method: 1) is able to predict the catheter distal response merely based on proximal input commands without including sensory feedback; 2) has a high modeling accuracy compared to other existing analytic models and machine learning models; 3) does not require a large amount of data and has a fast training procedure. The LSTM allows catheters to be steered precisely during endovascular interventions. A further contribution is that we use the developed model together with OCT-based distance sensing to reconstruct a 3D artificial environment. This chapter is structured as follows: Section 5.2.1 describes the structure of the LSTM as well as the experimental setup. Section 5.3 introduces the designed experiments to validate the LSTM. Section 5.5 concludes the work and proposes future work on catheter control.

5.2 Methods

5.2.1 Long Short-Term Memory network

A system is said to exhibit hysteresis if it has a sort of memory. This means that the output at a certain moment is determined not only by the corresponding input, but also by past inputs [256].

Table 5.1: Methods to minimize hysteresis in flexible surgical robots from previous literature

Authors and publication year	Device	Actuation/Transmission Principle	Methods
Bardou *et al.*, 2012 [245]	Endoscope	Cable-Driven (CD)	Feedback from Electromagnetic (EM) sensor
Cabras *et al.*, 2017 [248]	Endoscope	CD	Locating catheter tip using imaging modalities
Reilink *et al.*, 2013 [249]	Endoscope	CD	Locating catheter tip using imaging modalities
Baek *et al.*, 2020 [133]	Multi-arm Flexible Surgical Robot	Tendon-sheath mechanism (TSM)	Computer vision + kinematic model
Do *et al.*, 2014 [250]	Endoscope	TSM	Bouc–Wen model
Omisore *et al.*, 2018 [251]	Cardiovascular Catheter	CD	Adaptive Backlash Model
Legrand *et al.*, 2020 [252]	Fetoscope	Mckibben muscle	Generalized Prandtl-Ishlinskii Model
Xu *et al.*, 2017 [253]	Serpentine Manipulator	TSM	Machine learning (regression)
Porto *et al.*, 2019 [247]	Endoscope	CD	Machine learning + Inverse kinematic Model

LSTM was proposed as a tool to process sequential information and take historical information into account [79]. Therefore, it is not so illogical to consider the use of LSTM to model hysteretic behavior.

$$
\begin{aligned}
f^{(t)} &= \sigma(W_f p^{(t)} + U_f h^{(t-1)} + b_f) \\
i^{(t)} &= \sigma(W_i p^{(t)} + U_i h^{(t-1)} + b_i) \\
\widetilde{c}^{(t)} &= \tanh(W_g p^{(t)} + U_g h^{(t-1)} + b_g) \\
c^{(t)} &= f^{(t)} * c^{(t-1)} + i^{(t)} * \widetilde{c}^{(t)} \\
o^{(t)} &= \sigma(W_o p^{(t)} + U_o h^{(t-1)} + b_o) \\
h^{(t)} &= \tanh(c^{(t)}) * o^{(t)}
\end{aligned}
\tag{5.1}
$$

LSTM can be viewed as a stack of LSTM cells that contain a number of units in an LSTM cell (see Fig. 5.1). Each LSTM cell is equipped with an input gate, an output gate, and a forget gate. The information in a unit is processed following (5.1) [79]. In (5.1), $p^{(t)}$ is the input pressure to a Pneumatic Artificial Muscle (PAM) at time t. The variable σ is the sigmoid function with an output range of [0,1]. It regulates the percentage of information flows through each gate. Another activation function is the tanh, whose outputs range from [-1,1]. The forget gate f determines which memory in the previous cell state $c^{(t-1)}$ is retained. The input gate i and the input modulation gate $\widetilde{c}$ are combined to update the cell state. Subsequently, a new cell state $c^{(t)}$ is created. The cell state c stores selective information that the LSTM processes all along, thus a long-term memory in the LSTM can be established. The output gate o works as a filter. It regulates which part of the information from the new cell state $c^{(t)}$ is output and transferred to the hidden state $h^{(t)}$. The matrices W, U and b, with different subscripts, represent the weights and biases. These gates keep track of the dependencies in the input sequence.

5.2.2 Experimental setup

In order to collect data for training and testing the LSTM, a bench-top experimental setup was built. A schematic diagram of this setup is illustrated in Fig. 5.2. This setup contains a one degree-of-freedom (DOF) unidirectional catheter distal segment with an embedded PAM (Fig. 5.2). The catheter is fabricated out of Nitinol using metal laser cutting technology and has a diameter of 4.4 mm. The custom-made artificial muscle is used to actuate the catheter. By increasing the pressure, the muscle contracts and applies, through a steer cable that is attached off-centered at the catheter tip, a bending moment on the catheter tip. In order to accurately control the input pressure, the pressure is fed by an air supply, via a proportional

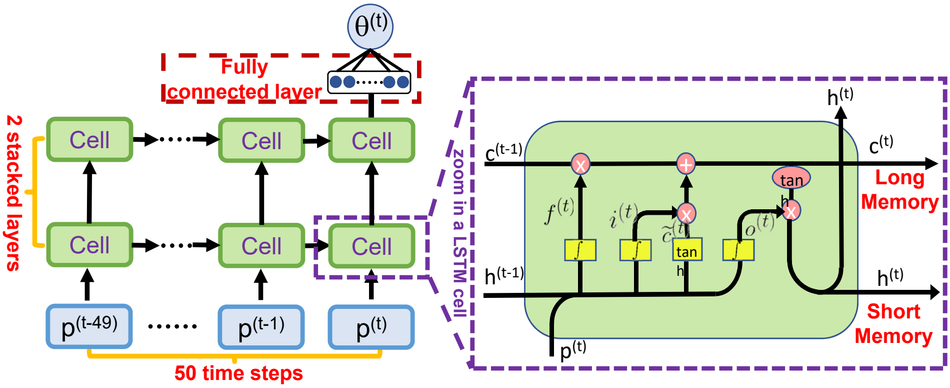


Figure 5.1: *Left*: a stacked 2-layer LSTM that consists of multiple LSTM cells, is unrolled into multiple time steps. In this work, the LSTM is comprised of 64 neurons per layer. A fully connected layer is added after the LSTM at time t. *Right*: each LSTM unit has three gates, i.e., input gate i, output gate o, and forget gate f that regulate the information flow according to (5.1). A window size of 50 is used to incorporate the input pressure p that was applied in the last 50 time steps. All this information is used to predict the future bending angle θ.

pressure valve (Festo, Germany) to the artificial muscle in cascade (Fig. 5.2). The proportional valve receives control signals from an analog output module NI-9263 (NI, Texas, USA). A pressure sensor (21Yseries, Keller, Switzerland) is installed in series with the above-mentioned circuit in order to measure the pressure value. A laser photoelectric sensor (OADM 12I6460/S35A, Baumer Group, Switzerland) is used to measure the distance s (unit: mm) from the sensor to the catheter tip with a sampling frequency of 250 Hz. A reflective sheet is attached on the catheter tip to increase the area of the reflective surface. The measured voltage output U (unit: V) from the laser sensor can be converted to the catheter tip bending angle θ (unit: deg) using following relation:

$$s_i = mU_i + n \tag{5.2}$$

$$d_i^{'} = \max(s_1, s_2, \cdots, s_N) - s_i \tag{5.3}$$

$$\theta_i = \frac{2 \cdot d_i^{'} \cdot L}{H^2 + d_i^{'2}} \cdot \frac{180}{\pi} \tag{5.4}$$

In (5.2) - (5.4), subscript i indicates the i-th sample in a group and the N represents the number of samples in a data group. The voltage U measured by the laser photoelectric sensor is proportional to the distance s to the measured object, thus $m = 4.05$ (mm/V) and $n = 29.36$ (mm) in (5.2) was identified only once prior

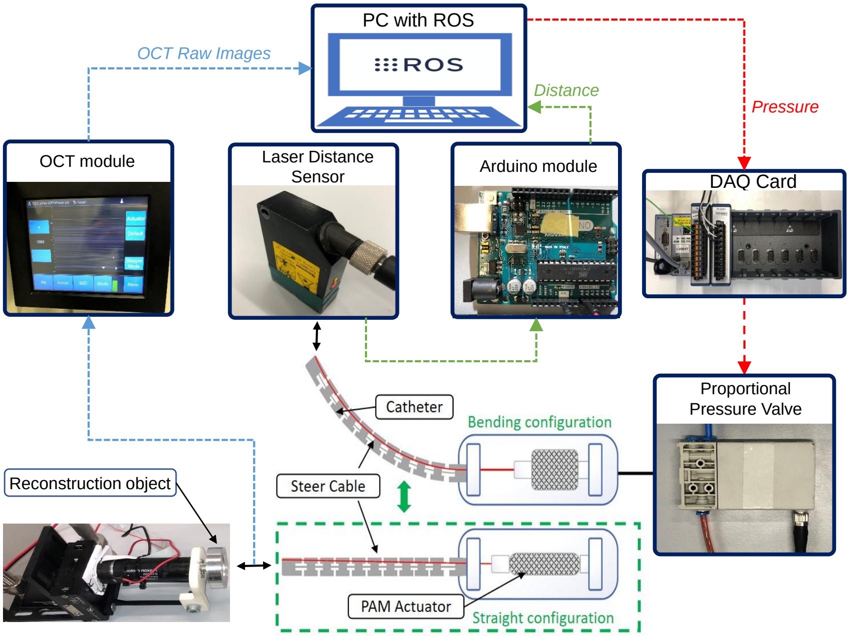


Figure 5.2: A PAM-driven catheter segment is controlled by a proportional pneumatic valve, which receives a command signal from a PC through a NI® CompactRIO system. The resulting catheter tip bending angle is captured by a laser sensor. A ROS-based GUI is created for pressure control and data collection.

to data collection. Referencing the straight configuration as zero displacement, then the displacement of the catheter tip $d^{'}$ can be calculated by subtracting the measured distance s from the maximum distance $\max(s_1, s_2, \cdots, s_N)$ between the catheter tip and the laser sensor as expressed in (5.3). Subsequently, as shown in Fig. 5.3, the catheter bending angle θ can be obtained based on a constant curvature model [225] assumption, in which the gravitational and inertial forces of the catheter segment could be assumed to be negligible. The data from the distance sensor are collected by an Arduino® microcontroller, and then sent back to the PC via a Robot Operating System (ROS) topic. A Graphical User Interface (GUI) is created in ROS to facilitate the users to switch among different excitation signals, which are detailed in Section 5.2.3.

5.2.3 Training data acquisition

In order to provide the LSTM with data containing diverse excitation patterns, descending sinusoidal pressures with zero baseline described in (5.5) and with

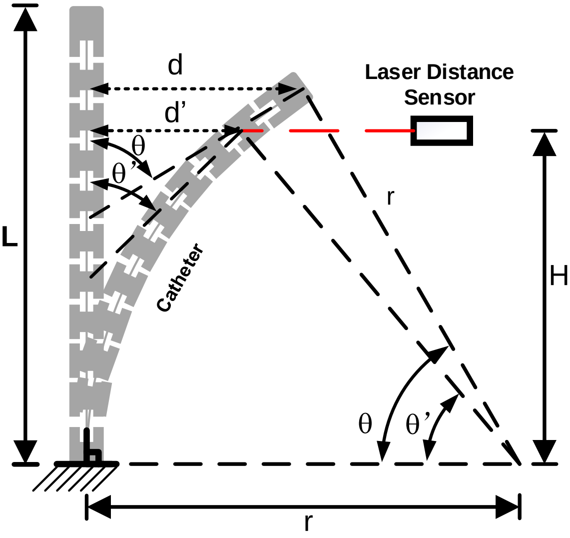


Figure 5.3: The catheter tip displacement $d^{'}$ captured by the laser distance sensor is converted to the bending angle θ using (5.2) - (5.4) based on a constant curvature model [225] assumption.

non-zero baseline described in (5.6) were sent to the setup to generate multi-loop hysteresis (Fig. 5.4).

$$p_1(t) = Ae^{-\tau t}(\sin(2\pi ft - \frac{\pi}{2}) + 1) \qquad [\text{bar}] \tag{5.5}$$

$$p_2(t) = Ae^{-\tau t}(\sin(2\pi ft - \frac{\pi}{2})) + A \qquad [\text{bar}] \tag{5.6}$$

The amplitude A of both signals is set at 1.5 bars to achieve a maximum amplitude of 3 bars. In (5.5) and (5.6), the variable f is the excitation frequency in Hz. Input frequencies up to 1 Hz are investigated in this work. The hysteresis behavior depends both on the input frequency and the amplitude of the excitation signal [257]. Different excitation frequencies are included in the training data, namely the frequencies of the excitation signal are set to 0.2, 0.4, 0.6 and 0.8 Hz in the training data set. The time constant τ was chosen with only one value $\tau = 0.15$ to generate multi-loop hysteresis.

Two types of training data (bending angle - pressure) featuring major and minor loops were obtained as shown in Fig. 5.4. In total, 26798 samples were acquired in the training data set. The training data reveals widening hysteresis loops with increasing excitation frequencies. One can also observe deadbands at the bottom of the loops (Fig. 5.4). It is noteworthy that the hysteresis does not only come

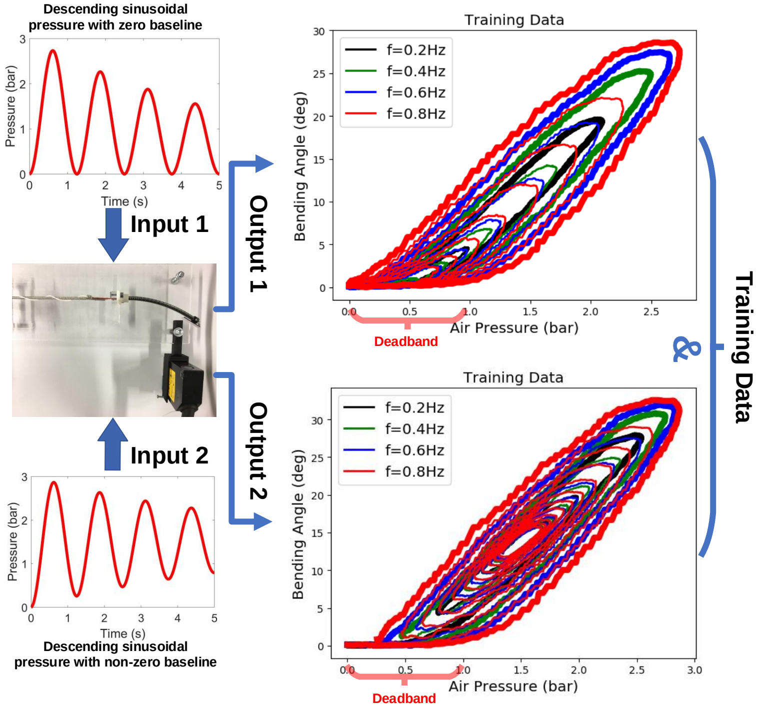


Figure 5.4: The training data are collected from the experimental setup: bending angle (deg) is visualized as a function of the input pressure (bar): *(Output1)*: 4 groups of data with zero baseline are collected based on combinations of $f = 0.2$, 0.4, 0.6, 0.8 and time constant $\tau = 0.15$ following (5.5); *(Output2)*: 4 groups of data with non-zero baseline are collected following (5.6) based on the same parameters as in output 1. Thick lines represent the major hysteresis loop with maximum excitation, while thinner ones indicate the minor loops.

from the PAM, but also due to friction from e.g. relative movements of the steer cable and NiTi tube during bending, the compressibility of air, the compliance of the pneumatic tubes, the nonlinear behavior of the employed material (Nitinol) and of the valves that were used. All these phenomena contribute to the resulting asymmetric, saturated, rate-dependent hysteresis behavior.

5.2.4 Hyperparameters of the LSTM and network training

A 2-layer stacked LSTM (see Fig. 5.1) was adopted. According to a previous pilot study [258], a window size equal to 50 was shown to have good performance while maintaining a low computational cost. Therefore, the input pressures were segmented into a window size of 50. In other words, sequential pressure data

Table 5.2: hyperparameters for the LSTM network

LSTM	Number of hidden layers	Number of neurons per cell	Activation functions	Optimizer	Loss function
	2	64, 64	Tanh/Sigmoid	Adam	L2 Loss
	Training-subset/ Validation ratio	Batch size	Learning Rate	Epoches	
	70%/30%	16	0.001	50	

$p^{(t-49)}$,$p^{(t-48)}$,...$p^{(t)}$ were entered into the LSTM to predict one bending angle $\theta^{(t)}$. The zeros were padded on the left of the input sequence to predict the first 49 bending angles. A fully connected layer (input dimensions = 64, output dimension =1) was added after the last LSTM cell. To increase the learning efficacy, all training data were normalized between [-1,1] to match the range of the activation function in (5.1). The hyperparameters of the LSTM are listed in Table 5.2. The LSTM is simple and fast with only two hidden layers and 64 neurons per layer. These two hyperparameters were chosen and tested empirically, as adding more layers or neurons not only led to worse fitting, but also increased the computational cost and the needed training time. The total number of trainable parameters is 50497. The network was implemented in Pytorch, an open-source machine learning framework. The training was performed using a 4 GB NVIDIA CUDA-capable GPU. The LSTM was trained for 50 epochs and the whole training time was around 10 minutes.

5.2.5 Modeling evaluation methods

Three metrics i.e. the Maximum Absolute Error (MAE), the Root Mean Square Error (RMSE) and the Normalized Root Mean Square Error (NRMSE) are used to quantitatively evaluate the model performance in this paper. The MAE that is computed following (5.7) measures the maximum absolute difference between the predicted bending angles and ground truth among all the samples:

$$MAE = \max\left\{\left|\hat{\theta}_i - \theta_i\right|\right\}, i = 1, 2, ..., N \tag{5.7}$$

N is the number of sample points in each group of test data. The RMSE following (5.8) calculates the square root of the square difference between the predicted bending angles and the ground truth.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{\theta}_i - \theta_i)^2}{N}}, i = 1, 2, ..., N \tag{5.8}$$

The NRMSE relates the RMSE to the observed range of the variables, and it is defined as:

$$NRMSE = \frac{RMSE}{\theta_{max} - \theta_{min}} \tag{5.9}$$

5.3 Experiments and results

5.3.1 Preliminary evaluation of the LSTM

In the preliminary evaluation, a well-trained LSTM was tested on four types of signals (see Fig. 5.5, 5.6, 5.7, 5.8). The LSTM, similar to other neural networks, consists of a huge amount of weights and biases that are distributed in each layer. The output of the LSTM is derived by multiplying the inputs with the embedded weights and biases. Therefore, the experiments in this section were merely carried out once, since an LSTM would predict the same results if the input remains the same.

A state-of-the-art analytic model called Deadband Rate-Dependent Prandtl-Ishlinskii (DRDPI) model proposed in [257] was established for comparison to the LSTM. The DRDPI model is a sophisticated and practical model that takes into account the impact of frequency on the hysteresis pattern. In addition, deadband operators in this model allow it to model asymmetric hysteresis as well as saturation-exhibiting behavior that appears at the bottom of the hysteresis loops (see Fig. 5.4). Therefore, the DRDPI model is fully adapted to model the hysteresis originating from the entire PAM-driven catheter system. The parameters of the DRDPI model were identified using a genetic optimization algorithm in MATLAB® Toolbox. The training data remained the same as for the LSTM (see Fig. 5.4). The identification process was performed on CPU (Intel Core i7 CPU @ 2.80GHz with a RAM of 8GB) as there was no wide-spread library for GPU-based training. The entire identification procedure of the DRDPI model took around 3.5 hours. A machine learning model, i.e. Support Vector Regression (SVR) was also implemented to enrich the comparison. SVR is an expansion of a classical Support Vector Machine (SVM) from pure classification to regression tasks. Similar to SVM, it is designed for estimation of high dimensional, nonlinear problems when only a limited number of samples are available [259]. The SVR model implemented in this paper was based on [259]. The hyperparameters of the SVR are chosen as follows: penalty C = 10, kernel = Radial Basis Function (RBF) kernel, kernel coefficient γ = 0.1, margin of tolerance ϵ = 0.1. Parameter C and γ were optimized using a grid search (C$\in\{0.001,0.01,0.1,1,10\}$, $\gamma\in\{0.001,0.01,0.1,1\}$) and a cross-validation method (k=10). The optimization results of the three models are shown in Table 5.3. The rightmost column represents the mean of the relative change in MSE loss

Table 5.3: optimization results of three models

Modeling Methods	Iterations	Training Loss (deg^2)	Mean of Relative Change
DRDPI	1000	0.344	0.013%
SVR	100	0.002782	0.708%
LSTM	50	0.00056	0.190%

in the last 10% of the iterations. The small mean of relative change indicates that the three models were adequately converged, and thus the comparison is fair.

The pattern of the test signals as well as the modeling results of the LSTM, the DRDPI model, and the SVR are shown in Fig. 5.5, 5.6, 5.7, 5.8 and described in detail as follows:

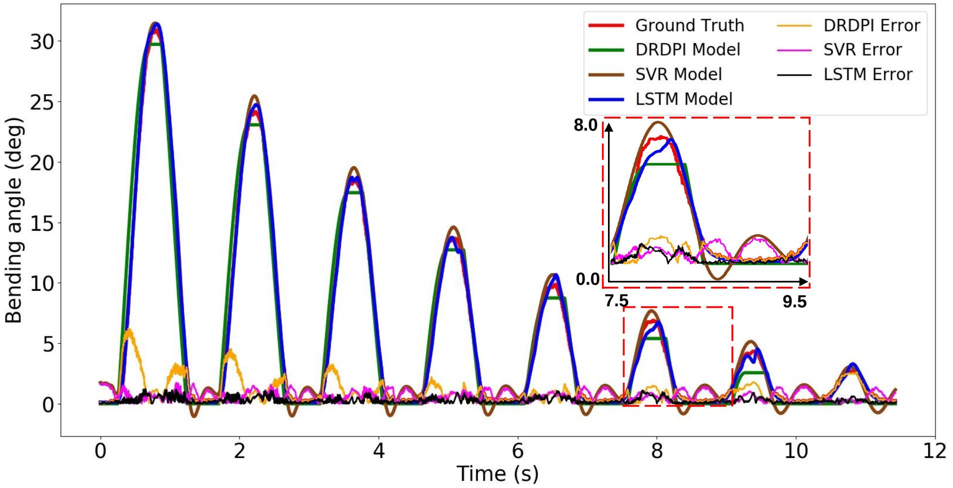


Figure 5.5: Descending sine wave differing from training data (f=0.7 Hz, τ=0.12 s^{-1})

a) A descending sine wave following (5.5) with f = 0.7 Hz and τ = 0.12 s^{-1} differing from the training data was generated, while the amplitude A = 1.5 bar remains the same.

The RMSE and the NRMSE of the LSTM are 0.36° and 1.17%, respectively. Both metrics outperform those achieved by the DRDPI model (1.40° and 4.54%) and the SVR model. In the first four loops, the error of the DRDPI tends to grow larger during the loading and unloading phase (see Fig. 5.5), whereas the MAE of 6.23° occurs in the first loop. Even though achieving a comparatively acceptable performance in few loops in the middle, the performance of the DRDPI model degrades again in the last loop. In contrast, the error of the LSTM consistently remains below 1.23° (MAE) throughout the range, which is even smaller than the

RMSE of the DRDPI model.

b) To test the performance of the LSTM on data with time-varying frequency, a pressure signal following (5.10) with A = 0.9 bar, B = 1.2 bar, f = 0. 7 Hz, τ = 0.1 s^{-1} and c = -0.1 Hz/s was utilized to generate attenuated down-chirp sine wave with non-zero baseline. The variable c is the chirpiness that regulates the rate of frequency change over time.

$$p(t) = Ae^{-\tau t}(\sin(2\pi(f+ct)t - \frac{\pi}{2}) + 1) + B \qquad [\text{bar}] \tag{5.10}$$

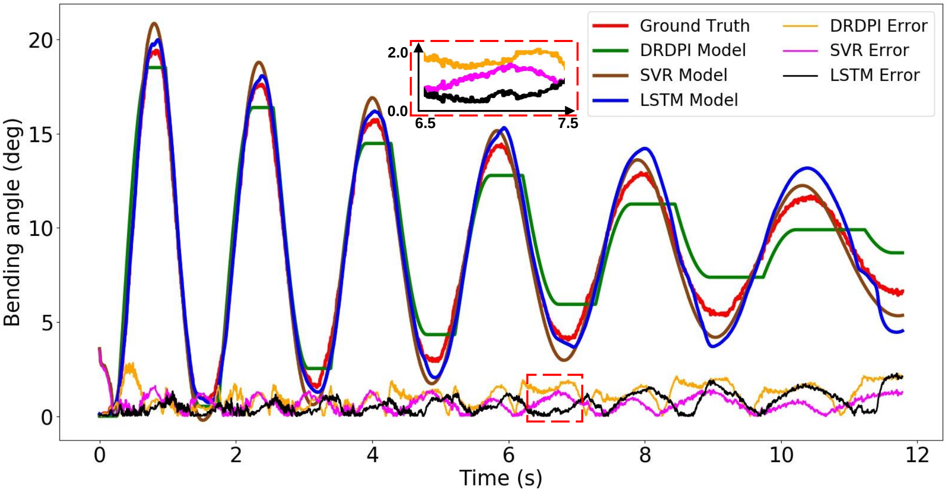


Figure 5.6: Attenuated down-chirp sine wave with shifted baseline (τ=0.1 s^{-1}, c=-0.1 Hz/s)

Unlike the DRDPI model that can merely model the saturating area by using a plateau, the LSTM and the SVR is able to predict a smooth sine-shaped curve when approaching extrema (see Fig. 5.6). This phenomenon can also be observed from other test signals. Consequently, it leads to lower RMSEs of the LSTM and the SVR with 0.82° and 0.80°, respectively, compared to DRDPI with 1.22°.

c) The third set of test data has an ascending pattern that is not contained in the training data. Following (5.11), with A = 0.6 bar, f = 0.3 Hz , τ = 0.05 s^{-1} and c = 0.15 Hz/s, an ascending up-chirp sine wave was generated.

$$p(t) = Ae^{-\tau t}(\sin(2\pi(f+ct)t - \frac{\pi}{2}) + 1) \qquad [\text{bar}] \tag{5.11}$$

The performance of the LSTM advances with increasing frequency, whereas the DRDPI model achieves a similar level of performance in the first two loops, before

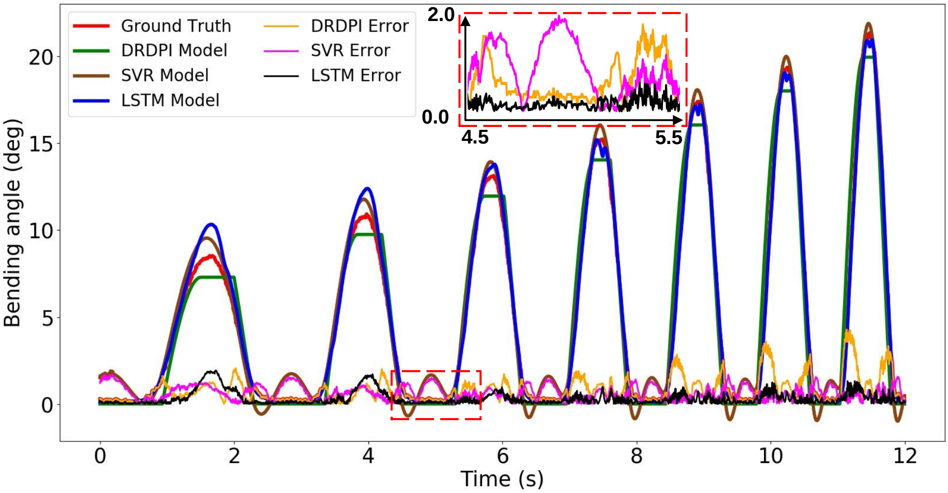


Figure 5.7: Ascending up-chirp sine wave with zero baseline (f=0.3 Hz, τ=0.05 s^{-1}, c=0.15 Hz/s)

it starts to deteriorate until reaching an MAE of 4.28° in the loading phase of the last loop (Fig. 5.7). Note that the SVR consistently predicts wavy lines in deadzones, causing its performance to be 34.2% worse than the LSTM in terms of the RMSE.

d) Arbitrarily varying signals represent the most general commands that can take place in practice. It is used to further explore the potential of the LSTM.

The LSTM consistently shows low errors, regardless of signal patterns, with RMSE and NRMSE of 0.59° and 2.42%, respectively. While the DRDPI also predicts a quasi-plateau, it can be seen that each plateau occurs at a considerable offset (Fig. 5.8), where its MAE of 2.98° emerges on the second plateau. Both the RMSE (0.59°) and the NRMSE (2.42%) of the LSTM are 60.1% and 36.0% smaller than the DRDPI model and the SVR, respectively.

The average inference time of the LSTM for a single point is approximately 25 ms based on the test data. Table 5.4 summarizes and compares the performance of the three models quantitatively based on the three metrics. In terms of NRMSE, apart from group (b) in which the LSTM is 33.1% better than the DRDPI model, the LSTM outperforms the DRDPI model by at least 50% in the other three groups. Moveover, except for achieving comparable performance in the ascending up-chirp sine wave, the LSTM is at least 34.2% better than the SVR. With the provided experiments, we showcased that the LSTM is capable of capturing dynamic responses of our system which are – given that our catheter is lightweight and the speeds are moderate – dominated by the valve dynamics. We believe that faster motion could also be captured if trained properly. Overall, the LSTM is

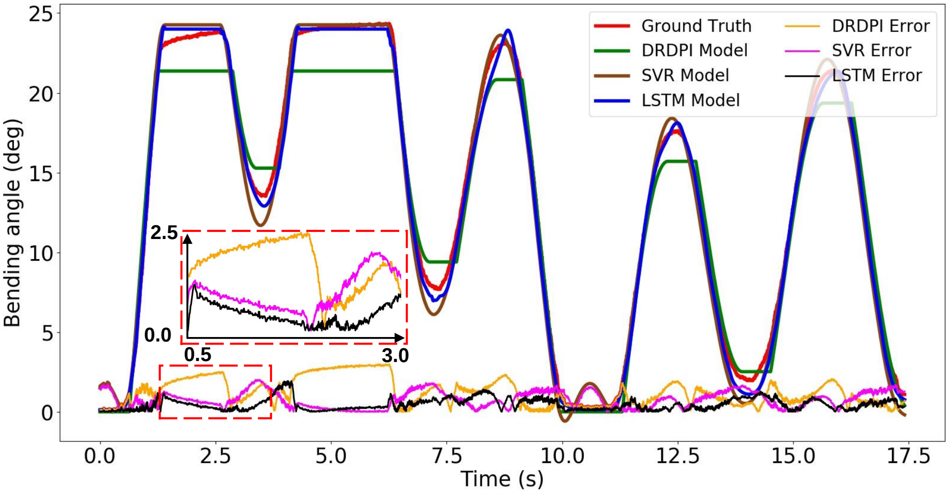
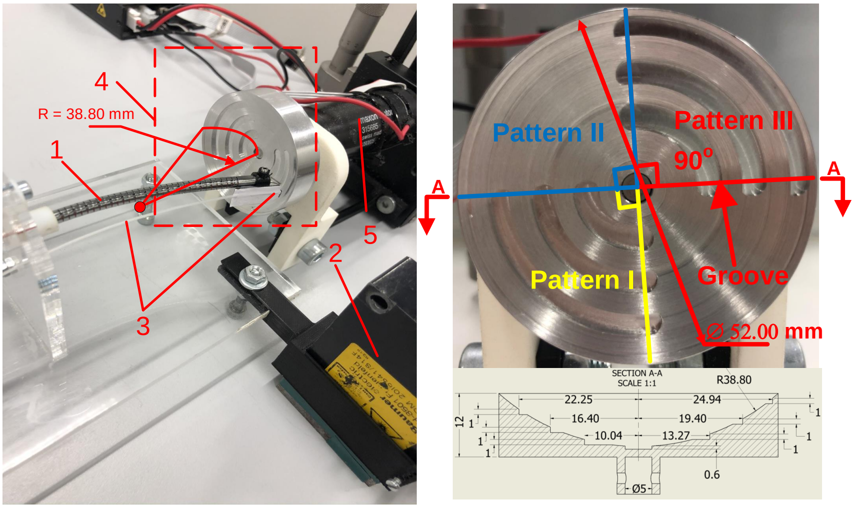


Figure 5.8: Arbitrarily varying signal with plateaus

Table 5.4: quantitative performance of the LSTM, the SVR and the DRDPI model on four groups of test sets.

Modeling Methods	Descending sine wave differing from training data			Attenuated down-chirp sine wave with shifted baseline		
	RMSE (deg)	MAE (deg)	NRMSE	RMSE (deg)	MAE (deg)	NRMSE
DRDPI	1.40	6.23	4.54%	1.22	2.83	6.28%
SVR	0.76	1.83	2.45%	0.80	3.54	4.10%
LSTM	0.36	1.23	1.17%	0.82	2.27	4.20%
Modeling Methods	Ascending up-chirp sinewave with zero baseline			Arbitrarily varyingsignal with plateaus		
	RMSE (deg)	MAE (deg)	NRMSE	RMSE (deg)	MAE (deg)	NRMSE
DRDPI	1.07	4.28	4.99%	1.48	2.98	6.07%
SVR	0.76	1.69	3.56%	0.92	2.01	3.78%
LSTM	0.50	1.92	2.34%	0.59	1.93	2.42%

able to accurately model asymmetric, saturated, and rate-dependent hysteresis originating from the entire catheter system with a sub-degree precision as well as a lead in performance compared to the DRDPI model and the SVR. In cardiovascular applications, the required precision that clinicians indicate as being acceptable is typically in the order of 1–3 mm [260], which corresponds to 2.09-6.26 degrees according to (5.4). Although confirmed in a simplified setting, the obtained performance of the LSTM (see Table 5.4) shows a good potential to satisfy the requirements in terms of precision also in a more realistic clinical setting.



(a) (b)

Figure 5.9: (a) overview of reconstruction module: 1. distal catheter segment; 2. laser distance sensor; 3. OCT fiber; 4. reconstruction object; 5. DC motor; (b) *(up)* three patterns distributed in a circular sector of 270° *(bottom)* cross-section view.

5.3.2 3D Reconstruction experiment

In a bifurcation lesion, angiography may not be able to accurately image the carina area because the main vessel and the side branch overlap each other. As a solution, a catheter-based forward-looking intravascular ultrasound (FL-IVUS) could help reconstruct the anatomy of the lesion [261]. In this case, the positioning accuracy of the catheter tip is very critical, otherwise a precise 3D reconstruction cannot be guaranteed. A 3D reconstruction experiment is depicted in this subsection (Fig. 5.9). A forward-looking optical coherence tomography (FL-OCT) fiber is used to simulate the A-mode FL-IVUS. The experiment is set up to test whether the modeling accuracy of the LSTM allows high-quality reconstruction of the environment. Considering that LSTM and SVR both rely on machine learning, while DRDPI is analytic, and given the superiority of LSTM over SVR, it was found to be more interesting to compare the performance of LSTM with the analytic DRDPI in the following experiment.

The experimental setup is shown in Fig. 5.2. Apart from the modules introduced in Section 5.2.2, a reconstruction module is added for this experiment. A hollow spherical cap with embedded grooves (Fig. 5.9) is fabricated to serve as a dummy environment. This cap is placed in front of the catheter tip. There are three patterns on this object, and each of them covers a circular sector of 90°. The

Table 5.5: the results of 3d reconstruction experiment (unit: mm, red: max. error, green: min. error)

Angle (°)		LSTM		DRDPI			LSTM		DRDPI			LSTM		DRDPI	
		Mean	Std	Mean	Std		Mean	Std	Mean	Std		Mean	Std	Mean	Std
0	Pattern I	0.15	0.16	0.54	0.46	Pattern II	**0.14**	0.16	0.53	0.42	Pattern III	0.24	0.29	0.69	0.57
15		0.18	0.21	0.57	0.46		**0.14**	0.17	**0.52**	0.41		0.26	0.30	**0.73**	0.59
30		**0.11**	0.13	**0.50**	0.41		**0.19**	0.21	0.54	0.45		0.24	0.29	**0.73**	0.58
45		**0.22**	0.25	**0.58**	0.47		0.15	0.19	0.53	0.43		0.26	0.32	0.67	0.56
60		0.18	0.21	0.51	0.44		0.18	0.21	**0.56**	0.43		**0.27**	0.31	0.68	0.57
75		0.15	0.21	0.56	0.42		**0.14**	0.17	0.54	0.41		**0.22**	0.28	**0.64**	0.54
90		0.14	0.18	0.52	0.42		0.17	0.19	0.55	0.45		0.24	0.29	0.69	0.57
Overall		**0.16**	**0.04**	**0.54**	**0.03**		**0.16**	**0.02**	**0.54**	**0.01**		**0.25**	**0.02**	**0.69**	**0.03**
p-value		**0.002117**					**0.002021**					**0.001998**			

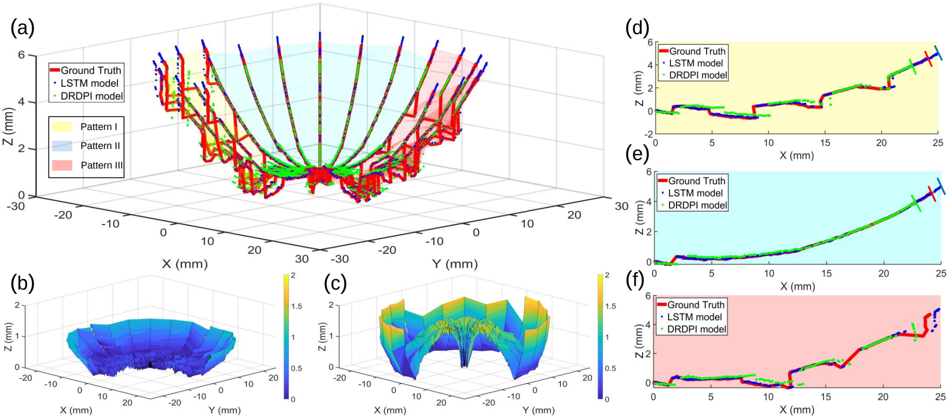


Figure 5.10: Results of 3D reconstruction experiment: (a) reconstruction of a hollow spherical-cap object with an OCT scan at every 15° using lateral distance sensor (ground truth, red), LSTM (blue dots) and DRDPI model (green dots); (b) - (c) reconstruction errors of LSTM and DRDPI model respectively represented in color-coded maps; (d) - (f) middle cross-section (at 45°) of each pattern showing the planar reconstruction performance of the two models.

dimensions of the object as well as the patterns can be seen in Fig. 5.9. Patterns I and III have three milled grooves at various distances with respect to the center axis, while pattern II is a smooth surface without any grooves. The object is rotated by a DC motor (Maxon Group, Switzerland) so as to simulate the rotation motion of the catheter. An OCT fiber (iOCT, OptoMedical Technologies GmbH, Germany) is attached onto the catheter tip to measure the front distance to the object by using an A-mode scan. Raw A-mode images acquired are published on a ROS topic and processed in Matlab (Fig. 5.2) using a method introduced in [262]. A harmonic sinusoidal pressure ($A = 2.95$ [bar], $f = 0.1$ Hz) is used to bend the catheter in order to cover the entire range of the semispherical cap. The scan object remains still during one back and forth motion of the catheter, and then rotates for 15° for the next scan until all the three patterns are scanned. The lateral distance sensor measures the actual tip bending angle. By combining the bending angle measured by the laser with the OCT measurement, followed by a conversion from Polar to Cartesian Coordinate, an approximation of the ground truth profile of the environment is obtained. This profile is then compared with the environment that is computed from the modeled bending angle (LSTM and DRDPI respectively) and the OCT distance measurement. If the hysteresis was perfectly compensated, both reconstructions should be matched. The reconstruction error is defined as the mean of the Euclidean distance between the ground truth and the reconstructed points derived by the two models.

The overall reconstruction performance is shown in Fig. 5.10(a). A quantitative evaluation can be found in Table 5.5. Figure 5.10(d) shows the 45 degree planar reconstruction of pattern I. One can observe that the LSTM can reconstruct the cross-section profile with high accuracy (reconstruction error = 0.22 mm, Table 5.5). Both the shape and the location of the grooves are well reconstructed except that the reconstructed curve is slightly stretched at the end with respect to the ground truth. On the contrary, the DRDPI model achieves an acceptable performance on curve, whereas the locations of the first and second grooves are reconstructed with a large variation, leading to a large reconstruction error of 0.58 mm (Table 5.5). In Fig. 5.10(e) and Fig. 5.10(f), the reconstruction performance of the two models follows a similar trend in Pattern I. The color-coded maps in Fig. 5.10(b) and Fig. 5.10(c) describe the reconstruction error across all the patterns. The reconstruction error of the LSTM gradually rises up until reaching around 1mm when approaching the edge of the object. Unlike the LSTM, the DRDPI model has a large reconstruction error around the center axis, as it utilizes deadband operators to model the dead zone of the hysteresis. After achieving its best performance in the halfway, its performance degrades again and the reconstruction error approaches ca. 2 mm. The reason for this is explained below. Unlike the LSTM, which can track the area near upper turning points with a smooth sine-shaped curve, the DRDPI model uses deadband operators i.e. plateaus to model the upper turning point which lead to a large modeling error. As shown in Table 5.5, the average reconstruction errors of the LSTM are 0.16 mm, 0.16 mm and 0.25 mm in pattern I to III. These errors are 70.4%, 70.4%, 63.8% lower than those of the DRDPI model. To ensure that the difference in results between the two models is not caused by sampling errors, a Mann-Whitney U Test is used to compare the LSTM and the DRDPI model. The null hypothesis H_0 that the mean reconstruction errors of the two models were not statistically different, could be rejected as a p-value smaller than 0.01 was found in all three patterns. The results prove that performance of the LSTM and the DRDPI model are statistically significantly different.

The reconstruction experiment confirms the feasibility that, with the assistance of the LSTM, a PAM-driven catheter can overcome the complex inherent hysteresis and accurately reconstruct the anatomical structure with forward-looking imaging modalities in intravascular interventions.

5.4 Extended exploration of hysteresis modeling

5.4.1 Introduction

In the previous sections, the hysteresis phenomenon in a pneumatically driven catheter was modeled using DL methods. However, this hysteresis was only 1-DOF *unidirectional*. The tip bending angle of the robotic catheter was determined using a laser distance sensor, which cannot be used in clinical scenarios. In this section, the DL method was further explored to understand whether it can: 1) be applied to another experimental setup with a different type of hysteresis; and 2) be used to model 1-DOF *bidirectional* hysteresis. In addition, another sensing modality, namely FBG sensor, was adopted to determine the catheter bending angle. Compared to the laser distance sensor, this measurement method is better suited for clinical scenarios as it can be readily integrated into the instrument and does not suffer from line-of-sight problems. It is worth noting that modeling 1-DOF bidirectional hysteresis is not as straightforward as replicating what has been accomplished in one direction. The deadzone surrounding the neutral position poses the hardest challenge to tackle. Therefore, new data collection and training techniques are anticipated. The work is conducted in the framework of the ARTERY project of which the background is explained below.

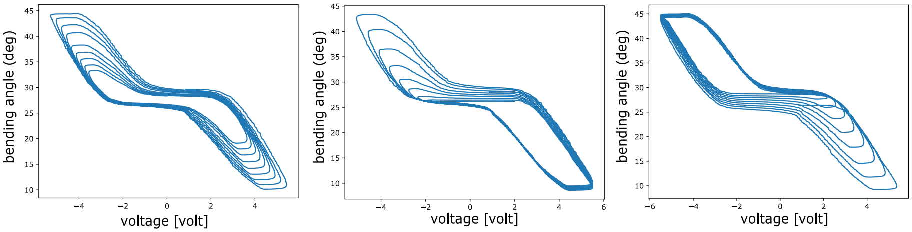
5.4.2 Clinical background

Mitral regurgitation is a cardiac condition characterized by a malfunctioning mitral valve in the heart, which is not able to close tightly and thereby allows blood to flow backward. A MitraClip system is commonly utilized to manage this condition [263]. During this procedure, the installed clip helps the mitral valve close more efficiently. However, physicians often face challenges operating the current MitraClip system manually. The system consists of a Steerable Guide Catheter (SGC) and a clip delivery catheter. The SGC in particular can be bent in one plane by operating a knob. The SGC is cable-driven and as such is prone to friction at the knobs and gears. Other problems include deadband (implying that the catheter tip does not respond to increases in air pressure and remains in its original configuration), backlash, and spring-back during use. These factors contribute to the highly non-linear nature of the system. Additionally, the setup is quite cumbersome, which could complicate the procedure.

In order to tackle these problems, the ARTERY project, funded by the EU (GA No 101017140), seeks to develop a robotic SGC. The robotic SGC is improved from the existing MitraClip system, enhanced by the integration of advanced actuators such as PAMs or motors. This design could mitigate the physical burden of physicians. However, whether manual or automated, hysteresis still persists. This problem negatively impacts the precise steering of the SGC. Therefore, the hysteresis must be understood and compensated for. This study proposes to use the LSTM to model all of the non-linear effects in the catheter system in one shot.

5.4.3 Data collection and ANN training

The training and testing datasets for the LSTM were collected with the pneumatically-operated Steerable Guide Catheter (SGC). The SGC system is actuated by means of PAMs that provide a contraction force when supplied with compressed air. The operating range of these PAMs is typically from 0 bar to 8



(a) (b) (c)

Figure 5.11: Datasets of three training groups with a pattern shown in each subfigure: (a) convergence to zero; (b) maximum positive amplitude; (c) minimum negative amplitude. The catheter presents an initial bending angle of approximately 27 degrees when it remains unactuated.

bar. The PAMs incorporated in this setup provide an axial contraction force of up to 100 N for a contraction length of 10 mm at 8 bar input pressure. The workspace of the SGC tip is ±30° from the initial angle which is approximately 27°.

$$V = Ae^{-\tau t}\sin(2\pi f t) \quad [\mathrm{V}] \tag{5.12}$$

$$V = Ae^{-\tau t}(\sin(2\pi f t) + 1) - A \quad [\mathrm{V}] \tag{5.13}$$

$$V = Ae^{-\tau t}(\sin(2\pi f t) - 1) + A \quad [\mathrm{V}] \tag{5.14}$$

Attenuated sinusoidal waves in voltage were employed to excite the system. The attenuated sinusoidal commands have three different patterns (see Fig. 5.11) that converge to 0, amplitude A, and amplitude $-A$, as described in (5.12), (5.13), and (5.14) respectively. Amplitude A was set to a value of 5.5 volts. Our purpose in utilizing this data collection strategy is to systematically record and encompass the broadest possible spectrum of catheter movements. Data were also collected at four different frequencies $f = 0.025$, 0.05, 0.1, and 0.2 Hz, along with two τ values: 0.0025 and 0.01 s^{-1}. The value of τ is used to regulate the descending speed. The combination of three trajectory patterns, four frequencies, and two τ values form 24 groups of datasets. These data were collected at a rate of 40 Hz. The total number of training data is 171163. The bending angle of the catheter is measured by the Fiber Bragg Grating (FBG) sensor. From the shape of the FBG sensor that follows the bending of the catheter, two vectors are obtained that represent the tip of the bendable section and the base of the bendable section. Each vector is selected by choosing two consecutive points from the curvature point cloud. The spacing between the two vectors from this point-cloud is based on the known spacing between the the gratings of the FBG ssensor, as well as the known length of the bendable section of the catheter [217]. Once the two vectors are obtained, the bending angle in one plane can be calculated from the vectors $\vec{p_b}$ representing the base position vector and $\vec{p_t}$ representing the tip position vector as follows:

$$\theta_m = \cos^{-1}(\vec{p_b} \cdot \vec{p_t}) \tag{5.15}$$

The training set was divided into a training subset and a validation set, maintaining a ratio of 70% to 30%. An LSTM, designed with four layers, each containing 128 neurons, was used. A fully connected layer was appended after the last LSTM layer. All input and output data were normalized in the range of [-1,1]. For weights initialization, the Xavier uniform initialization method [264] was employed for the input layers, while Orthogonal initialization was utilized for the hidden layers. A learning rate of 0.005 was adopted, with a decay factor of 0.2 implemented after every 20 epochs. Adam was the optimizer of choice, and the Mean Squared Error (MSE) was utilized as the loss function. The LSTM was trained for 80 epochs. The code was implemented using PyTorch. GPU was used to accelerate the training procedure. RMSE, MAE and NRMSE [239] were used to evaluate the performance of the LSTM.

5.4.4 Results and discussions

The performance of the LSTM was verified using an ascending sinusoidal trajectory. The results displayed an RMSE of 0.54 degrees and an NRMSE of 1.17%. In addition, an MAE of 1.94 degrees occurred within the deadzone region. There is an evident reduction in the performance of the LSTM throughout the initial five minor loops, correlating with the deadzone roughly spanning from 25 to 30 degrees. This observation is further highlighted in the pressure-bending angle diagram. Figure 5.13 illustrates the performance of the LSTM in response to an arbitrarily varying voltage. Areas experiencing an error greater than 2 degrees are colored orange. The RMSE and MAE are 1.22 degrees and 3.51 degrees, respectively, both of which surpass the numbers associated with actuation using ascending commands. It is noteworthy that all orange areas fall within the dead zone region (approximately 25 to 30 degrees), suggesting that the prediction of sheath steering is most challenging in this working range. Additionally, the color-coded subfigure (b) in Fig. 5.13 highlights that large errors predominantly occur when the direction of catheter movement was changed.

5.4.5 Conclusion on the extended exploration of hysteresis modeling

Section 5.4 introduces a data-driven approach to model the bidirectional rate-dependent hysteresis of a steerable guide catheter. The trained model was validated

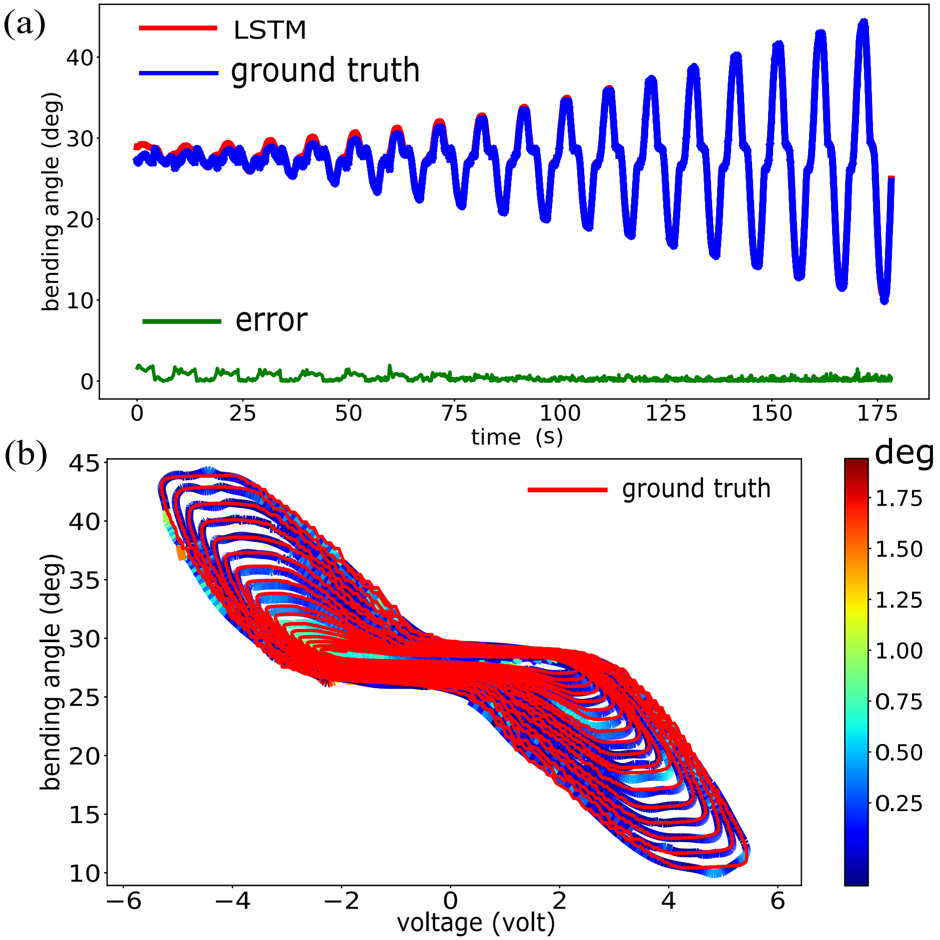


Figure 5.12: Test results when the system is actuated using an ascending control commands: (a) the red line signifies the ground truth, while the blue line represents the LSTM's predictions. The prediction error is indicated by the green line; (b) the bending angle is represented as a function of voltage. A color-coded map displays the magnitude of error.

on two trajectories that feature patterns distinct from those in the training data (see Fig. 5.11). The experiment demonstrates that the RMSE is 1.22 degrees, which corresponds to 2.64% of the operational range, when assessed on a randomly varying trajectory. This demonstrates a comparable performance to the results shown in Section 5.3. Precise modeling allows the catheter tip to be accurately located without installing a sensor, which is a challenging endeavor due to space constraints. In addition, installing sensors at the tip could potentially impede

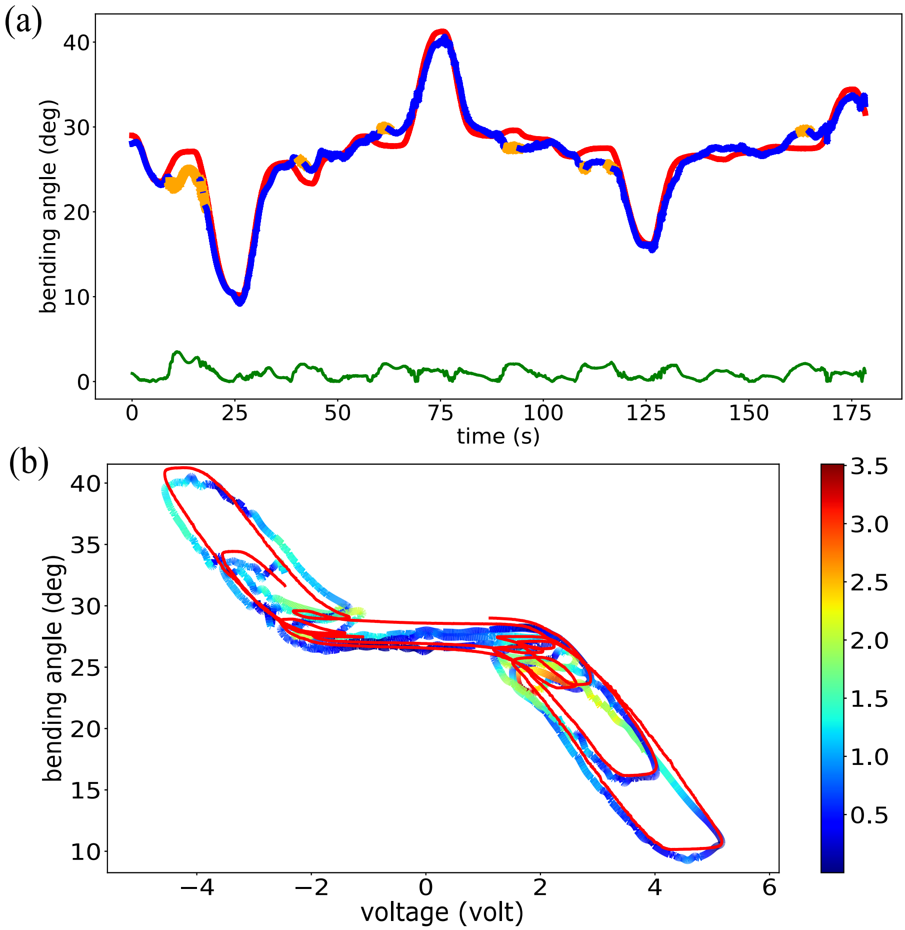


Figure 5.13: Test results when the system is actuated using arbitrarily varying commands. The legend remains consistent with that presented in Fig. 5.12. Additionally, the color orange in the (a) denotes areas where the error exceeds 2 degrees.

catheter functionality. For future work, it is envisioned that a deep-learning-based controller, capable of compensating for hysteresis and accurately steering the catheter in free space, will be developed.

5.5 Conclusion and future work regarding the DL-based hysteresis modeling

To address the hysteresis problem in robotic catheters (sub-objective 4), an LSTM network is proposed. An experimental setup containing a one-DOF PAM-driven catheter segment was developed for model validation. Descending sinusoidal pressure signals under four frequencies were used to generate multi-loop hysteresis, which was used as training data for the LSTM. The acquired data reveal a complex pattern, namely asymmetric, saturated, rate-dependent hysteresis, which originates from the entire catheter system rather than from the PAM itself. A DRDPI model and a SVR model were established for comparison with the LSTM.

The LSTM was first validated on four test signals containing diverse patterns, e.g. chirp signals, arbitrarily varying signals. The RMSE and MAE of the LSTM remain below 0.82°, 2.27°, respectively, across different test signals. The results reveal that the LSTM performs better than the DRDPI model and the SVR. Next, the LSTM was further validated through an OCT reconstruction experiment. The results indicate that the LSTM, with its good modeling accuracy, is able to reconstruct the object with errors below 0.25 mm in all three patterns, while the smallest reconstruction error achieved by the DRDPI model is 0.54 mm. Additionally, the effectiveness of the LSTM was validated on another device with 1 DOF bidirectional bending capability. In this validation, the bending angle was measured using an FBG sensor, while the LSTM continued to demonstrate comparable performance.

In general, LSTM demonstrates its robustness when the catheter system exhibits complex hysteresis. LSTM requires a similar amount of training data compared to analytical models and machine learning methods, and is very fast to characterize. The LSTM reduces the need to mount a sensor at the distal tip for localization, which is challenging due to spatial limitations and sterilization requirements. The proposed LSTM makes it possible to steer robotic catheters with good precision. This ability was demonstrated in a free space scenario (representing motion in the heart), which could enable interesting clinical applications such as detailed intra-operative reconstruction of anatomic lumens or cavities. The limitation of this chapter lies in the fact that validation has only been conducted on devices with 1-DOF bending ability, without verifying its performance in the 2-DOF scenario. Second, the trained LSTM is only applicable in free space, while the question of how to generalize the model to constrained space remains an area worth exploring.

In this chapter, only modeling of the hysteresis has been focused on. However, to improve the accuracy of catheter tip positioning, compensation for hysteresis is required. Additionally, the emphasis in this chapter has been limited to the position control of the catheter, without consideration given to the interaction forces between the catheter and its surrounding environment. In the next chapter, compensation for hysteresis is introduced. The effectiveness of the DL-based controller is validated through trajectory-following experiments. Furthermore, it is demonstrated that, based on the DL-based position controller, the contact forces between the catheter and its environment can be effectively minimized, avoiding the need for end-of-vessel sensors.

Chapter 6

Deep-learning-based hysteresis compensation and compliant motion control of robotic catheters

This chapter is based on the following publication:

D. Wu*, X. T. Ha*, Y. Zhang, M. Ourak, G. Borghesan, K. Niu, F. Trauzettel, J. Dankelman, A. Menciassi and E. Vander Poorten,“Deep-learning-based Compliant Motion Control of a Pneumatically-driven Robotic Catheter”, *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8853-8860, 2022

In cardiovascular interventions, when steering catheters and especially robotic catheters, great care should be paid to prevent applying too large forces on the vessel walls, as this could dislodge calcifications, induce scars or even cause perforation. To address this challenge, this paper presents a novel compliant motion control algorithm that relies solely on position sensing of the catheter tip and knowledge of the catheter's behavior. The proposed algorithm features a data-driven tip position controller. The controller is trained based on a so-called *control Long Short-Term Memory Network* (control-LSTM). Trajectory following experiments on four different trajectories are conducted to validate the quality of the proposed control-LSTM. The performance was compared with the performance of a controller that makes use of an analytical hysteresis model, i.e. the inverse Deadband Rate-Dependent Prandtl-Ishlinskii (IDRDPI) model. The results demonstrated superior positioning capability with sub-degree precision of the new approach in the presence of severe rate-dependent hysteresis. Experiments in a simplified setup as well as in an aortic phantom further show that the proposed approach allows reducing the interaction forces with the environment by around 70%. This work shows how deep learning can be exploited advantageously to avoid tedious modeling that would be needed to precisely steer continuum robots in constrained environments such as the patient's vasculature.

This chapter aligns with Sub-objective 5, focusing on elevating the level of autonomy in robotic catheters through a compliant motion control method and demonstrating this advancement in an *in-vitro* environment. This could potentially reduce the physical burden on clinicians and increase catheter control precision, which is often affected by the varying levels of experience among physicians. Ultimately, the method developed in this chapter could lead to safer and faster interventions.

6.1 Introduction

The modeling and actively steering of robotic catheters remains a challenge. This complexity comes from the fact that the catheter configuration depends both on steering maneuvers and on a complex distributed interaction with the environment. Managing the interaction force between the catheter and surrounding anatomy to an acceptable level is crucial to avoid tissue damage. Especially the interaction at the tip is important, as this is the most acute part of the catheter that can easily pierce the tissue or dislodge calcification.

If the configuration, i.e. the pose and/or shape, of the catheter's steerable segments could be measured, such information could be used advantageously to control the catheter so that it could be steered in a more gentle fashion through a vessel. Several methods have been proposed in the literature to achieve this goal. Kesner *et al.* presented a method in which a force sensor was integrated at the tip of a robotic catheter to keep the tip interaction force below potentially damaging forces [265]. This method requires a dedicated force sensor, which limits the functionalization of the catheter as there will be limited room left for integrating other functions. In addition, the structural properties of the catheter may be negatively affected by such a force sensor. Other approaches make use of complicated nonlinear mechanical models to achieve compliant motion control [266], [267]. Said models are typically very specific to the robotic structure at hand. Furthermore, significant modeling efforts are needed to obtain a very precise characterization of the input-output behavior of a specific continuum robot. To circumvent complex and time-consuming modeling, model-less control methods have been investigated in the past. A Multiagent Deep Q Network (MADQN), based on reinforcement learning, was employed by Ji *et al.* to control a cable-driven continuum surgical robot [268]. Yip *et al.* proposed a model-less controller based on a real-time estimated Jacobian of the continuum manipulator [269]. Follow-up work by Yip *et al.* further developed a hybrid position/force controller enabling safe interactions with the unknown constrained environment [270]. However, this model-less method relies on a dedicated distal force sensor and requires excellent accuracy from that sensor. No matter the approach, whether model-based or model-less, additional complexity stems from the non-ideal behavior of the actuator and the transmission system, where phenomena such as friction, hysteresis, or backlash present additional challenges to controlling the continuum robots [271, 272]. Reliable compliant motion control, therefore, requires accounting for both robot kinematics and the said inherent non-linear behavior. Compared to previous work, the proposed method aims to learn all these aspects in one shot.

Hysteresis can be compensated by closed-loop control using electromagnetic (EM) sensing [273], optical trackers [274] or medical imaging [275]. Alternatively, hysteresis can also be compensated in open-loop using feed-forward analytical models [134, 276, 277, 278]. Open-loop approaches are appealing, as they eliminate the need to install sensors at the distal end. However, creating reliable analytic models often requires tedious identification procedures. Next, if possible, the identified analytical model somehow needs to be inverted.

Deep learning (DL), albeit a black-box method, has gained growing attention due to its ability to accurately represent complex nonlinear behavior. Traditional Machine Learning (ML) or DL methods have been investigated in the past to model continuum robots. In [271], a neural network was developed to estimate the cable tension of a tendon-driven continuum robot in unloaded motion. Porto *et al.* proposed to use an Extreme Learning Machine, a supervised learning method, to model the inverse kinematics of a flexible endoscope and realise 3D position control [247]. Three data-driven approaches were proposed by Xu *et al.* to learn the inverse kinematics of a flexible surgical manipulator [253]. These approaches were validated by a trajectory tracking experiment. Among three approaches, the K-nearest neighbors regression achieved the lowest Root Mean Square Error (RMSE) of 2.1275 mm. Both studies [247, 253] focused on cable-driven robots, and adopted traditional ML methods e.g. regression, but did not investigate the benefit of the latest advances in DL. Our previous research [239] used an LSTM to *model* the hysteresis of a catheter system with sub-degree precision of the tip bending angle. On the other hand, this chapter proposes an open-loop controller based on a dedicated LSTM, which is different from the modeling LSTM of [239]. The LSTM is used to *control and compensate* for the hysteresis in a catheter system. In order to distinguish it from the LSTM of [239], the LSTM in this chapter is referred to as *control-LSTM*. The control-LSTM, which basically is a free-space position controller, forms the basis for the proposed compliant motion control algorithm. In practice, the LSTM-based controller will react to external forces by moving the catheter tip position to a pose that reduces the force acting at the tip. In summary, the main contributions of this work are:

- an open-loop controller based on an LSTM to control and compensate for the hysteresis in a catheter system. The proposed control-LSTM is used as a feedforward free-space catheter controller robust to severe hysteresis. It is precise and has a simple training procedure. Moreover, the control-LSTM is a direct inverse mode. In other words, the control-LSTM does not require a separate inversion step to be used in a controller, as would be needed by most traditional analytical hysteresis models, such as [139].
- a compliant motion controller based on the developed control-LSTM position

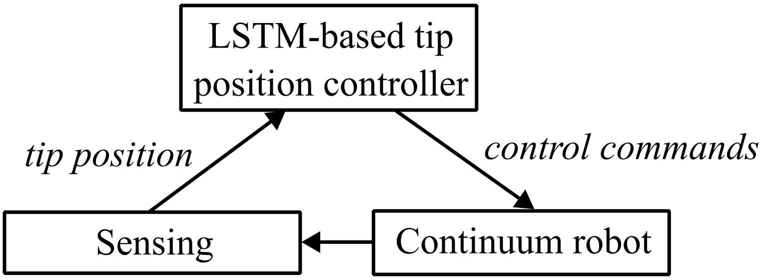


Figure 6.1: The block diagram of the proposed LSTM-based compliant motion controller.

controller that uses pose sensing as input without needing force- or contact-sensing and without sensing actuator torques.

- validation of the proposed controller both in a simplified setup as well as in an aortic phantom.

The rest of the chapter is organized as follows: Section 6.2 introduces the concept of DL-based compliant motion control. Section 6.3 details the experimental setup, the collection of training data, and the structure of the control-LSTM. Section 6.4 describes the trajectory following experiments. Next, the validation of the compliant motion control is displayed in Section 6.5 and Section 6.6. Section 6.7 concludes the chapter.

6.2 DL-based compliant motion catheter control

A new DL method that enables compliant motion control, and at the same time capturing the non-ideal behavior of the employed actuation, is introduced in this section.

The objective of our proposed compliant motion controller is to advance the robot's tip to the position provided by the sensing block, and thus minimize the force applied to the tip. To overcome the problem of hysteresis in catheters driven by cable or fluidics, instead of using the traditional kinematic model-based method [279] to control the catheter tip position, a control-LSTM position controller (hereinafter referred as "LSTM controller") was implemented to control the catheter tip position. The control-LSTM can predict the actuator commands that can advance the catheter tip to a conformed position, where the interaction force is minimized. This approach helps reduce the effort to model complex non-linear behavior and simplifies the identification procedure of catheters. Moreover, the proposed LSTM compliant motion controller requires only one position tracking sensor to measure the tip position. The block diagram of the proposed LSTM-based compliant motion controller is shown in Fig. 6.1.

To prove the concept, the proposed control-LSTMcompliant motion algorithm was implemented on a 2-DOF catheter segment (see Fig. 6.3) actuated by PAMs in one direction. The details of control-LSTM training and validation are described in Section 6.3 and Section 6.4. The overall block diagram of the LSTM compliant motion controller is shown in Fig. 6.2. Note that the proposed methods can be generalized to other sensors that are able to provide real-time estimates of the catheter's tip pose such as fluoroscopy [211], electrical impedance tomography [213], bending resistance [214] or Fiber Bragg Grating-based shape sensing method [229]. However, in this work, the EM tracker serves as a practical sensor to track tip position in order to enable compliant motion control.

6.3 Experimental setup and LSTM training

6.3.1 Experimental setup for validating the control-LSTM

This work uses a catheter system actuated by PAMs. However, PAM-driven catheter systems suffer from rate-dependent hysteresis, as shown in Chapter 5.

A bench-top setup (see Fig. 6.3) has been built for experimental validation. The setup contains a 85 mm long distal catheter segment with a diameter of 4.4 mm. The segment consists of a Nitinol notched-tube backbone. These notches make the catheter bendable in 2 DOFs [200]. In this work, the algorithm was validated in 1 DOF, while expanding this algorithm to 2-DOF will be part of our future work. A PAM is connected off-center by a steel cable to the catheter tip. By inflating the PAM, the PAM contracts. The resulting cable tension exerts a torque on the catheter tip, making it bend sideways. A proportional pressure valve (Festo Corporation, Germany) is employed to regulate the pressure in the PAM. A laser distance sensor (OADM 12I6460/S35A, Baumer Group, Switzerland) or an EM tracker (Northern Digital Inc., Canada) is used to measure catheter tip motion. For more details on the catheter system, please refer to [239]. Figure 6.2(b) (bottom, right) shows the rate-dependent, saturated, and asymmetric hysteresis of the PAM-driven catheter. The deadband pressure is about 0.7 bar. The rate-dependent hysteresis offsets range from ca. 0.3 bar to 1.2 bar, as the frequency increases from 0.2 Hz to 0.8 Hz.

6.3.2 Training data acquisition

The training data were collected when the catheter bends in 1 DOF. To generate abundant training data, descending sinusoidal pressure patterns with zero baseline

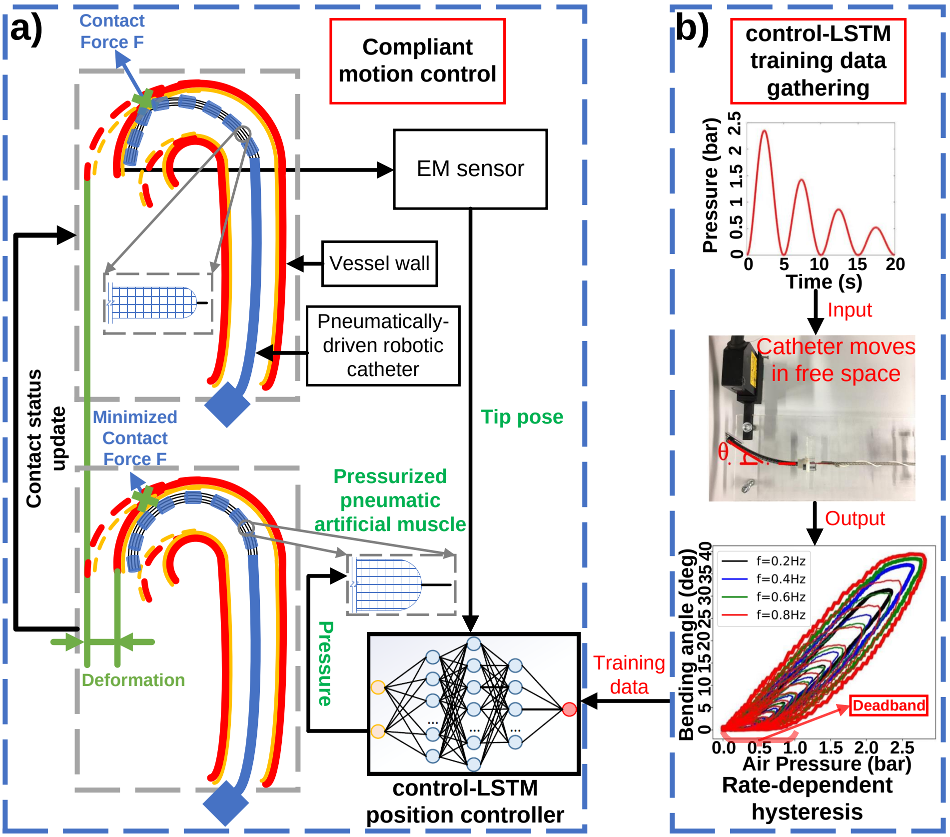


Figure 6.2: Overview of the compliant motion control algorithm: (a) Only tip pose information was acquired and then was input into the control-LSTM. The pressure predicted by the control-LSTM advances the catheter to a proper position where the interaction force is minimized. (b) Training data for the control-LSTM were collected during catheter motion in free space.

following (6.1) and with non-zero baseline, along (6.2), were sent to the catheter.

$$p_1(t) = Ae^{-\tau t}(\sin(2\pi f t - \frac{\pi}{2}) + 1) \qquad [\text{bar}] \tag{6.1}$$

$$p_2(t) = Ae^{-\tau t}(\sin(2\pi f t - \frac{\pi}{2})) + A \qquad [\text{bar}] \tag{6.2}$$

These signals generated the multi-loop hysteresis (Fig. 6.2(b)). To achieve a maximum pressure of 3 bars, the amplitude A is set to 1.5 bar. The variable f represents the excitation frequency (Hz). The time constant τ controls the rate at which the pressure increases or falls. This work only investigates the case where the frequency is less than 1 Hz. Figure 6.2(b) demonstrates that the formation of

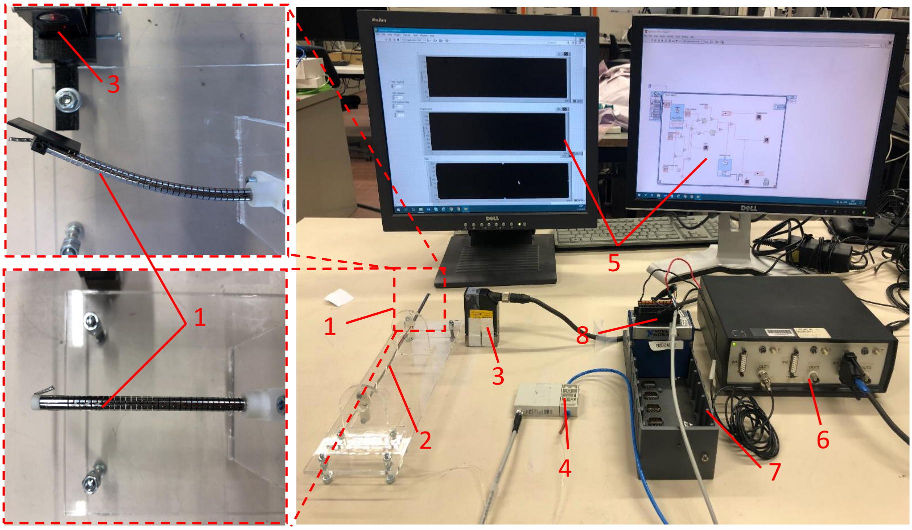


Figure 6.3: A bench-top setup for experimental validation: 1. catheter distal segment made of Notinol, 2. Pneumatic Artificial Muscle, 3. laser distance sensor, 4.proportional pressure valve, 5. Graphical User Interface (GUI) based on LabVIEW®, 6. control box, 7. embeded chassis, 8. NI® module. A close-up left up and left down show respectively the bending and straight configuration of the catheter segment.

hysteresis is affected not only by the frequency but also by the pressure amplitude. Therefore, different frequencies and pressure amplitudes are combined to excite the catheter system. The frequency f was switched between 0.2, 0.4, 0.6, 0.8, while τ was chosen as 0.02, 0.05, 0.1, 0.15, 0.2. Combining these two variables, as well as the signals from (6.1) and (6.2) resulted in 40 groups of training data. Measurements were sampled at 250 Hz (laser sensor), producing a training data set of 260269 samples in total.

In training, the bending angle is an input of the control-LSTM. Pressure, the parameter to be learned, was the output. Note that the control-LSTM's learning process is not straightforward. This is because the training data that is collected are somewhat idealized. During training, the input data (the bending angle) are noisy (due to measurement sensor noise) and exhibit hysteretic behavior, while the outputs are "clean" perfect sinusoidal pressure setpoints. Conversely, in practice, the control-LSTM will typically be provided with "clean" bending angle motion commands and is asked to predict the associated hysteretic pressure setpoints that would produce such smooth variations in the bending angle. For this to work, the control-LSTM thus needs to be able to generalize and learn itself the patterns that do not exist in the training data.

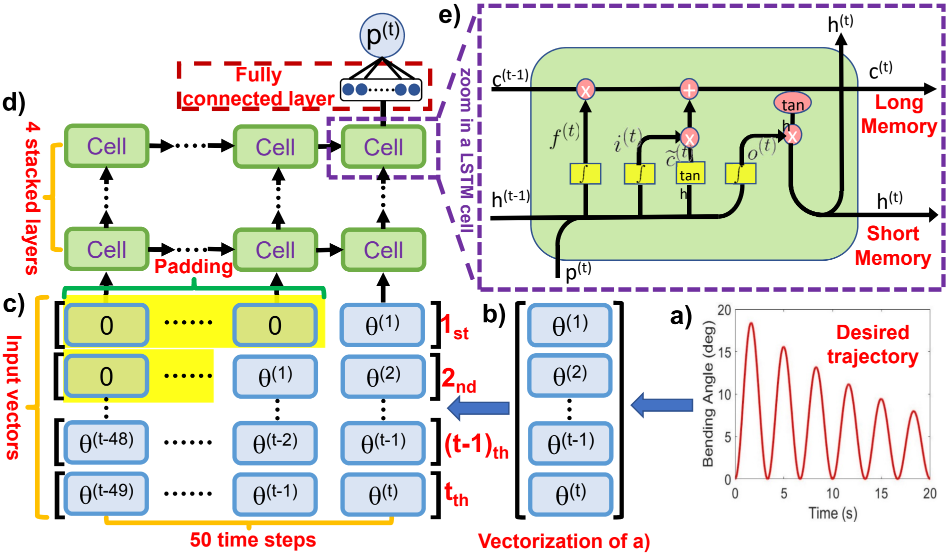


Figure 6.4: Inputs and structure of the control-LSTM (a)-(c): the trajectory is transformed into 50 × 1 long vectors that serve as input for the control-LSTM; (d)-(e) a stacked 4-layer control-LSTM that consists of LSTM cells, is unrolled into multiple time steps.

Table 6.1: Hyperparameters for the control-LSTM network

	Number of hidden layers	Number of neurons per cell	Activation functions	Optimizer	Loss function
control-LSTM	4	128	Tanh/Sigmoid	Adam	L2 Loss
	Training-subset/Validation ratio	Batch size	Learning Rate	Epoches	
	70%/30%	16	0.001	100	

6.3.3 Structure and hyperparameters of the control-LSTM

When a system exhibits hysteresis, the output of the system is not only determined by the current input, but also by historical inputs [280]. The ability of LSTM to learn historical information and use prior knowledge to predict system behavior at future time steps [281] formed a key motivation to use this network for this application. An LSTM can be viewed as a stack of LSTM cells (Fig. 6.4(d)), where each cell contains a number of units (Fig. 6.4(e)). Each LSTM cell is equipped with three gates i.e. an input gate i, an output gate o, and a forget gate f (Fig. 6.4(e)) that regulate the information flow [281].

Within this work, a 4-layer stacked control-LSTM (Fig. 6.4(d)) was used. The number of neurons per layer was set to 128. The pre-processing flow for the inputs into the control-LSTM is described as follows: a pre-defined trajectory is presented as a vector (Fig. 6.4(b)). The vector is split into multiple vectors with a length of 50 by means of a sliding window approach (Fig. 6.4(c)) as the inputs to the control-LSTM. The window size of 50 was found to balance well between model precision and computational cost [239]. As shown in the yellow highlighted area in Fig. 6.4(c), zeros were padded to the left for the first 49 input vectors, when there are not enough historical values present yet. A fully connected layer (input dimensions = 128, output dimension =1) was added after the last LSTM cell. All training data was normalized between [-1,1] to avoid large weights and to speed up the training procedure. The remaining hyperparameters of the control-LSTM are listed in Table 6.1. The network was implemented using PyTorch. The training procedure was carried out on a 4 GB NVIDIA® CUDA-capable GPU. The control-LSTM was trained for 100 epochs until the control-LSTM was adequately optimized. The whole training duration took approximately 2.25 hours.

6.3.4 Implementation of the IDRDPI model

To assess the performance of the control-LSTM, an inverse Deadband Rate-Dependent Prandtl-Ishlinskii (IDRDPI) model, introduced in [139], was implemented for comparison[1]. The IDRDPI model can compensate for hysteresis at different frequencies and also takes into account both major and minor loops. In addition, deadband operators in this model allow it to deal with the deadzone that appears at the bottom of the hysteresis loops (see Fig. 6.2(b)). Therefore, the IDRDPI model is well suited to tackle rate-dependent, saturated, and asymmetric hysteresis exhibited in the aforementioned catheter system. To identify the IDRDPI model, an inverse model of the Generalized Prandtl-Ishlinskii (GPI) was used, which provided a fairly accurate output for hysteresis compensation at a single frequency. The identification of the IDRDPI model in our work was divided into two steps: In the first step, four GPI models were identified by the hysteresis data at four frequencies i.e. 0.2, 0.4, 0.6, 0.8 Hz (19635 data points in total), then four corresponding inverse GPI models could be obtained by directly inverting the previously obtained GPI model. The inverse GPI models were able to generate training data for the IDRDPI model. In the second step, the identified inverse GPI model generated 150000 virtual training data, which was used to train the IDRDPI model. The parameters of the IDRDPI model were identified based on a genetic algorithm in MATLAB® Toolbox. The identification process was performed on a CPU (Intel Corei7-7700 CPU @ 2.80GHz with a RAM of 8.00GB) as there was no widespread library for GPU-based training. The whole identification procedure of the IDRDPI model took around 3.5 hours. The IDRDPI model was trained for 300 iterations. The mean of the relative change in the last 20% of iterations was just 0.07% when the stop condition was triggered, which revealed that the model was fully converged. As shown in Fig. 6.5, the IDRDPI model was also validated using the same procedure as the control-LSTM.

---

[1] The author would like to thank Saeid Shakiba for his helpful advice on the DRDPI model.

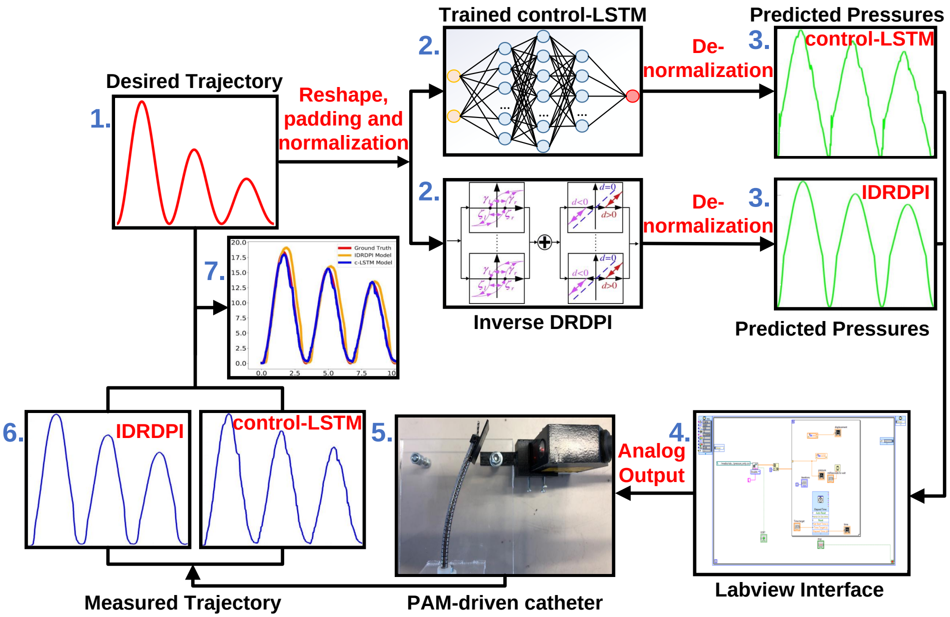


Figure 6.5: Experimental procedure to verify the performance of the LSTM and the IDRDPI model. 1) a desired trajectory is designed by the user; 2) the desired trajectory is pre-processed and fed into the LSTM and the IDRDPI model; 3) two models predict the corresponding pressures; 4) predicted pressures are read by Labview and applied to the catheter as feedforward control; 5) the catheter tip's motion is measured by a laser sensor; 6) measured voltages are converted to catheter tip bending angle; 7) both measured and desired trajectories are visualized and compared.

6.4 Validation of the LSTM-based controller

To evaluate the trained control-LSTM, an experimental procedure shown in Fig. 6.5 was conducted. Four types of trajectories that contain diverse patterns are tested. Three evaluation metrics i.e. the Maximum Absolute Error (MAE), the Root Mean Square Error (RMSE) and the Normalized Root Mean Square Error (NRMSE) were used to evaluate the performance of the LSTM controllers [239]. For each test trajectory, the validation procedure shown in Fig. 6.5 was repeated five times. The $\overline{MAE}$, $\overline{RMSE}$, and $\overline{NRMSE}$ represent the average value of the three metrics across the five attempts. The standard deviations of the three metrics were also calculated to assess the repeatability of this experiment. The designed test trajectories as well as the control results of both the control-LSTM and the IDRDPI model have been summarized in Fig. 6.6 to 6.10 and will be described next.

6.4.1 Response on descending sinusoidal trajectory

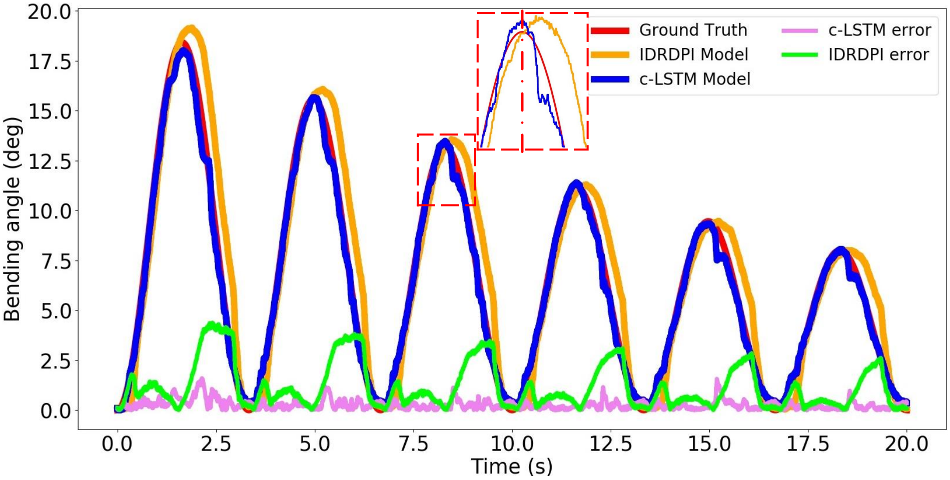


Figure 6.6: Descending sinusoidal trajectory (f=0.3, τ=0.05)

A descending sinusoidal trajectory (6.3) with $f = 0.3$ and $\tau = 0.05$ is generated as:

$$\theta(t) = Ae^{-\tau t}(\sin(2\pi ft - \frac{\pi}{2}) + 1) \qquad [\text{deg}] \tag{6.3}$$

with amplitude A set to 10. In terms of RMSE and NRMSE, the control-LSTM (0.36° and 1.95%) achieves a better performance than the IDRDPI model (1.69° and 9.17%). The MAE of the control-LSTM (1.60°) occurs in the unloading phase immediately after the peak. Here, the control-LSTM drops too fast after the peaks. This can be explained by the fact that the control-LSTM over-compensates for the dead zone at the top of the hysteresis curve. The IDRDPI model produces large offsets with respect to the set trajectory in the unloading phase. Figure 6.7 plots the actual bending angle versus the desired bending angle. A perfect compensation would lead to a straight line under 45 degrees. The figure shows how the LSTM controller produces a quasi-linear relation, whereas the IDRDPI model compensates for the hysteresis to a lesser extent.

6.4.2 Response on time-varying frequency trajectory

The performance of the control-LSTM on a trajectory with time-varying frequency is worth investigating. A pressure signal following (6.4) with $A = 10$, $B = 8$, $f = 0.5$, $\tau = 0.05$, and $c = -0.015$ was utilized to generate an attenuated down-chirp sinusoidal trajectory with shifted baseline:

$$\theta(t) = Ae^{-\tau t}(\sin(2\pi(f + ct)t - \frac{\pi}{2}) + 1) + B \qquad [\text{deg}] \tag{6.4}$$

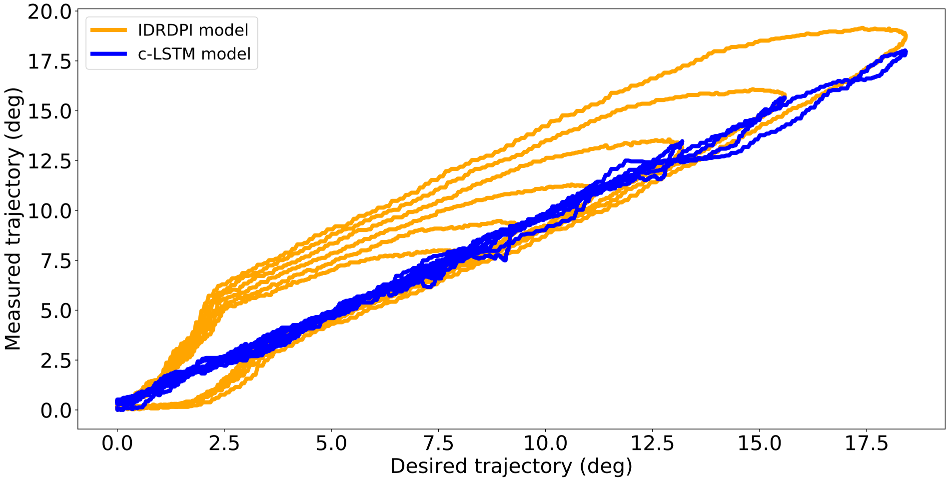


Figure 6.7: Relation between the desired trajectory (the same as in Fig. 6.6) and the measured trajectory: c-LSTM-based feedforward controller provides an identity mapping between desired trajectory and measured trajectory after compensation. However, the DRDPI model-based controller does not fully compensate for the hysteresis.

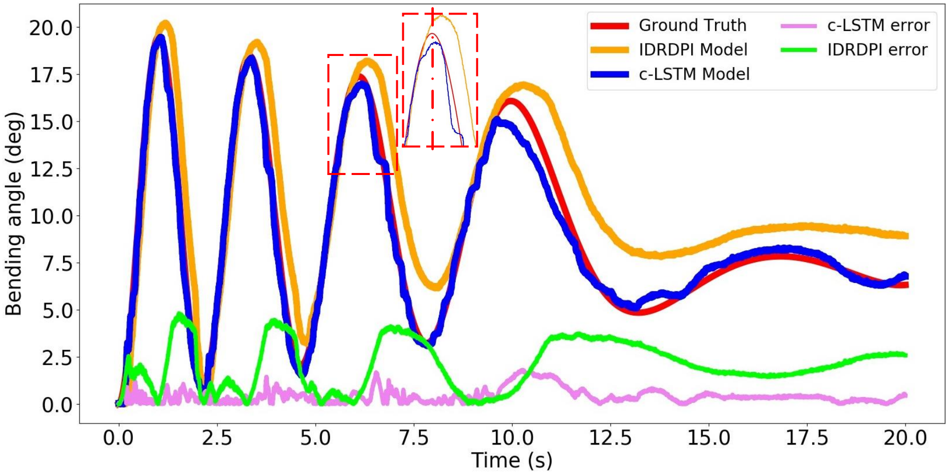


Figure 6.8: Attenuated down-chirp sinusoidal trajectory with shifted baseline (f=0.5, τ=0.05, c=-0.015)

The error of the IDRDPI fluctuates periodically (Fig. 6.8). Compared to the control-LSTM, the IDRDPI fails to accurately predict the peak and under-compensates in the unloading phase. This leads to the large RMSE (2.46°) and MAE (4.89°) of the IDRDPI. The error of control-LSTM remains low, achieving an RMSE of 0.58°. However, the error grows as the frequency decreases. Following this trend, the error of the control-LSTM reaches an MAE of 1.81° at the end.

6.4.3 Response on period-ascending trajectory

The third test trajectory is a period-ascending trajectory. This pattern was also not contained in the training data. Following (6.5), with $A{=}6$, $f{=}0.3$, $\tau{=}-0.02$ and $c{=}0.005$, an ascending up-chirp sinusoidal trajectory is generated through:

$$\theta(t) = Ae^{-\tau t}(\sin(2\pi(f+ct)t-\frac{\pi}{2})+1) \qquad [\text{deg}] \tag{6.5}$$

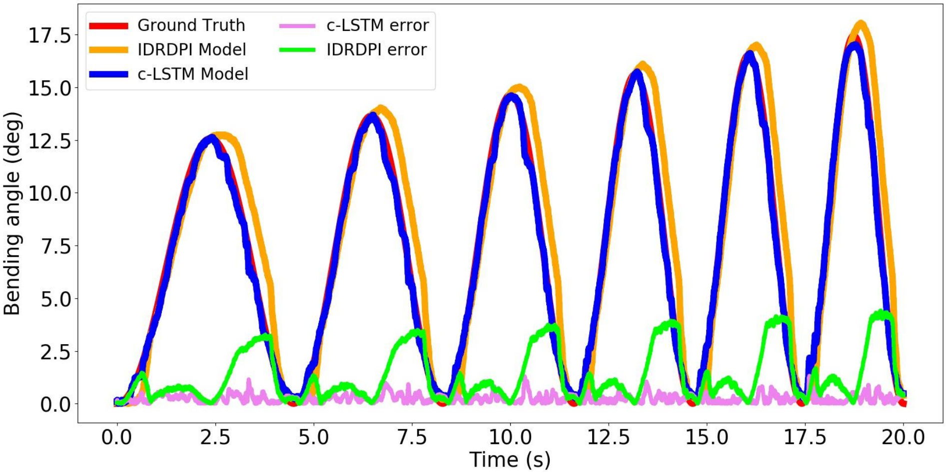


Figure 6.9: Ascending up-chirp sinusoidal trajectory with zero baseline (f=0.3, τ=-0.02, c=0.005)

As with the previous two tests, the IDRDPI model performs poorly in the unloading phase (Fig. 6.9). In contrast, the NRMSE and MAE of the control-LSTM are 1.82% and 1.32°, respectively, 82.8% and 70.5% better than the IDRDPI in terms of these two metrics.

6.4.4 Response on arbitrary trajectories

In practice, trajectories are likely to vary in a random fashion. Thus, arbitrarily varying trajectories, which are composed of multiple sinusoidal trajectories with distinct frequencies and amplitudes, have been tested to explore the versatility of the control-LSTM. In addition, some high-frequency small-amplitude variations are added. Such variations could, for example, arise if one wants to compensate for vessel deformations induced by heartbeat or blood flow. In such a case, to avoid contact with vessel walls, the catheter may need to perform low-amplitude high-frequency movements.

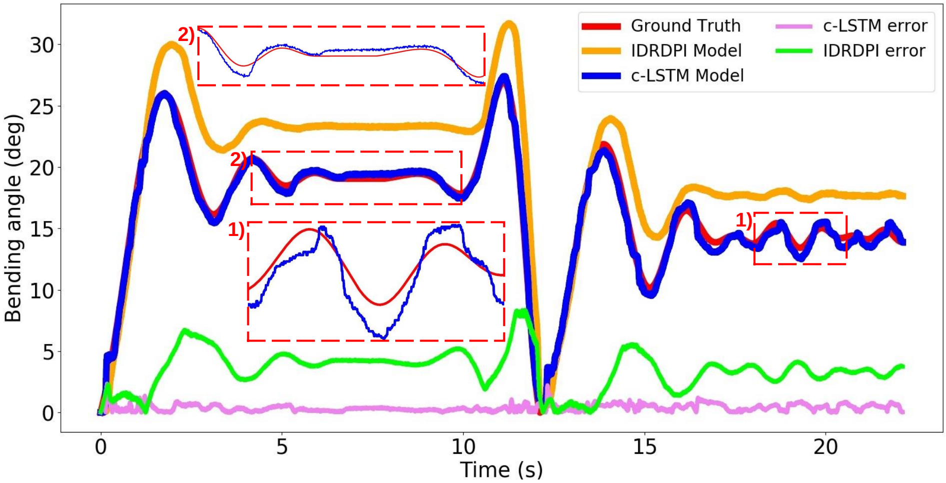


Figure 6.10: Arbitrarily varying trajectory with small variations, which are composed of multiple sinusoidal trajectories with distinct frequencies and amplitudes.

In general, the error of the control-LSTM is consistently very small, also in regions with small fluctuations (1st close-up in Fig. 6.10). It is worth noting that the control-LSTM is able to follow the set trajectory with high precision in the left plateau region as well (2nd close-up in Fig. 6.10). These confirm the small RMSE of 0.46° and MAE of 2.41° achieved by the control-LSTM. Contrary to this, the IDRDPI shows its worst performance among the four tested trajectories. The IDRDPI model consistently generates an offset from the set trajectory. Moreover, in the region of small variations, the IDRDPI model does not respond to small fluctuations like the control-LSTM, but rather produces a plateau.

6.4.5 Discussion of results of control-LSTM validation

The average inference time of the control-LSTM for a single output point is approximately 25 ms, thus it could be used in a real-time controller. Figure 6.11 summarizes and compares the performance of both controllers quantitatively according to the three metrics. Also, the standard deviation is given. In terms of RMSE, among the four test trajectories, both the control-LSTM and the IDRDPI

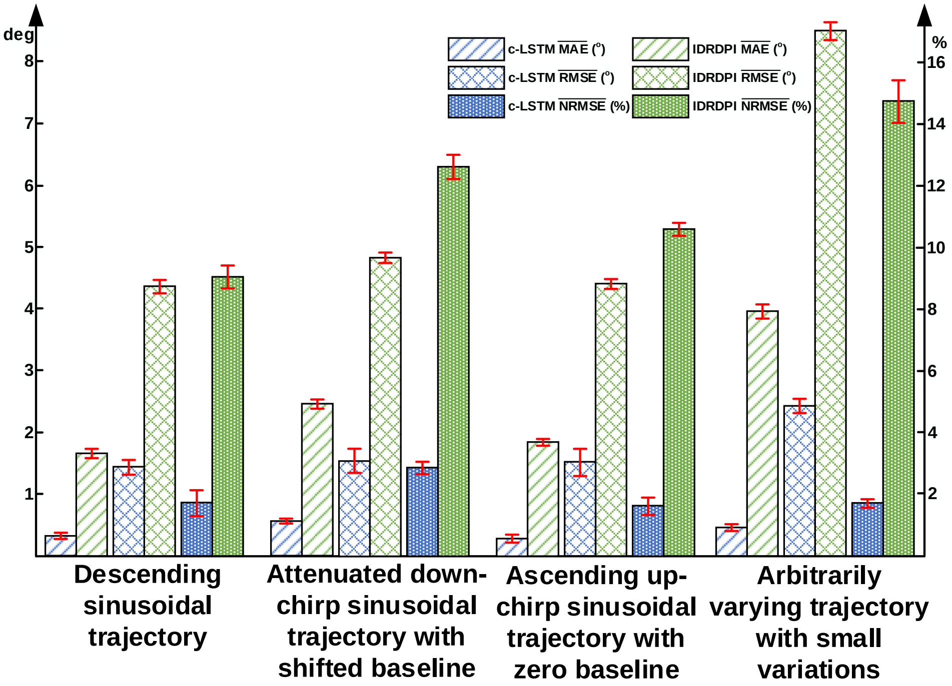


Figure 6.11: The $\overline{MAE}$ (°), $\overline{RMSE}$ (°) and $\overline{NRMSE}$ (%) of the c-LSTM (blue) and the IDRDPI model (green) on four trajectories.

perform worst on the arbitrarily varying trajectory. While the best performing trajectory for the control-LSTM is the ascending up-chirp sinusoidal trajectory. For the IDRDPI, the best results are obtained for the descending sinusoidal trajectory. In terms of NRMSE, the control-LSTM outperforms the IDRDPI model by 80.9%, 77.2%, 84.5%, and 88.3% in the four trajectories. The small standard deviations represented by the error bars in Fig. 6.11 reveal high repeatability across experiments. In cardiovascular interventions, the acceptable precision range is generally $1-3$mm [282], albeit sometimes specific scenarios may require better precision. This $1-3$ mm (in displacement) can be converted to $2.09^\circ - 6.26^\circ$ (in bending angle) using the formula introduced in [239], so in this regard, the controller based on the control-LSTM can meet the clinical requirements in terms of precision. This precise LSTM controller opens up new opportunities for lowering interaction forces as is shown in the next section.

6.5 Bench-top validation of the compliant motion controller

6.5.1 Experimental setup for validating the compliant controller

A steerable catheter (Fig. 6.12(a)) was used to demonstrate the compliant motion control capability. Compared to the catheter introduced in Section 6.3.1, a 6 mm diameter catheter was used to facilitate contact with a force sensor. In addition to diameter, structure and working principle remained the same as the catheter in Section 6.3.1. As shown in Fig. 6.12(a), an F/T sensor (Nano17, ATI Industrial Automation, USA) attached to an XYZ linear translation stage was used to push the catheter tip in one direction, simulating unknown interaction forces with a vessel or heart wall. During the experiments, the contact force of the F/T sensor and the displacement of the catheter tip measured by the EM sensor in the tip were recorded.

6.5.2 Design of the experiments

Two experiments were carried out. In the first experiment, contact was established at the catheter tip. In the second experiment, contact took place at half the length of the steerable segment. Each experiment began with the catheter in a straight configuration. The catheter was pushed by the F/T sensor until the muscle reached maximal pressure (5 Bar) to avoid contact force. The displacement was then reversed. These experiments were conducted first without and then with the compliant motion controller.

In the conducted experiments, the LSTM-based position controller, presented in Section 6.3.3, was used. The position controller ran at 40 Hz. Unlike in Section 6.4, where the bending angle was used as input for the LSTM controller, here the tip displacement d_t was used, as this metric is easier to calculate if an EM sensor was used. The tip displacement represents the Euclidean distance between the catheter's tip when unloaded, $P_{Neutral}$, to the current tip position, $P_{Current}$ (see Fig. 6.12(a)). Note that displacement is related to the bending angle, but it is a value that is more convenient to derive from the EM sensor data.

The displacement was used as input to the LSTM-based position controller to estimate the appropriate pressure needed to pressurize the catheter to the current position (Fig. 6.2) so that the interaction force between the catheter and the environment can be reduced. Forces were measured at 1 kHz, but then sub-sampled to allow for synchronization with the EM displacement data. Synchronization was done offline.

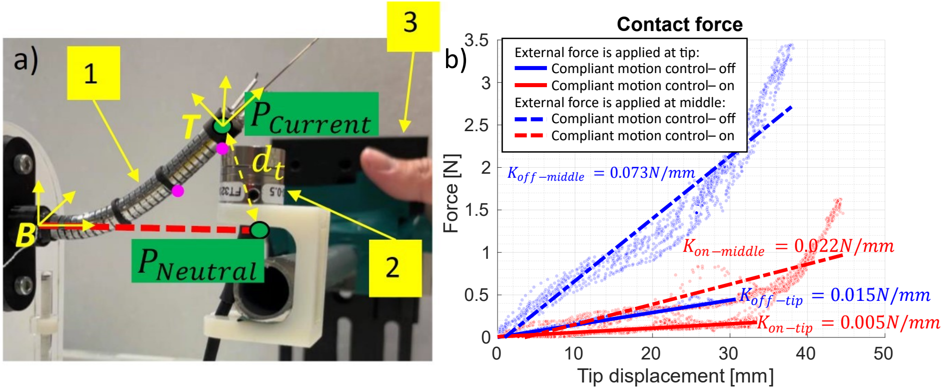


Figure 6.12: (a) Experimental setup to validate the compliant motion controller: 1. fluidics-driven catheter; 2. F/T sensor; 3. XYZ linear translation stage. The magenta circle shown the contact position in two experiments. (b) Measured force versus displacement of the catheter tip in case external force is applied at the tip (dashed line) and at the middle (solid line) of the catheter.

6.5.3 Results of the bench-top validation of complaint motion control

Figure 6.12(b) show the measured force versus the tip displacement when an external force is applied at the tip and half along the catheter length. The blue lines show the measured forces when the compliant motion controller was not active. The red lines depict forces when the controller was operational. To evaluate the efficiency of the proposed compliant motion controller, the perceived tip stiffness without (K_{Off}) and with (K_{On}) compliant motion controller were compared. Tip stiffness is measured as the ratio of the applied force versus the tip displacement. When the compliant controller is active, the tip stiffness was reduced by approximately 66% from 0.015 N/mm to 0.005 N/mm. Similarly, tip stiffness decreased from 0.073 N/mm to 0.022 N/mm (around 70%) when the external force acted on the catheter body.

6.5.4 Discussion of compliant motion control results

The complex relation revealed in Fig. 6.12(b) can be attributed to the fact that the force sensor slipped somewhat along the catheter length during the experiments. Although zero-force was targeted, the interaction force in Fig. 6.12(b) did not go down all the way to zero. The following reasons are expected to cause this: 1) in fact, as the catheter bends, the force applied by the F/T sensor will get both a tangential and a radial component. Only the radial force component is actually minimized by the here described compliant motion controller; 2) the control-LSTM based controller is not flawless, errors on that controller will cause some deviation as well; 3) the employed valve has limited bandwidth as well, so it can not react and instantaneously compensate an external force.

Nevertheless, the conducted experiments show that the proposed LSTM-based compliant motion controller can help reduce the force acting on the catheter at different places along the catheter length. Reductions of approximately 70% can be observed. This proves that deriving highly complex models that describe the mechanical behavior of the catheter is in fact not necessary for these kinds of tasks. Moreover, the proposed approach requires only one sensor to measure the catheter tip pose compared to other more complex methods [266, 267, 271, 283]. Even though the algorithm was only validated in a single bending plane, considering that the catheter can rotate around its axis and has identical mechanical properties in each bending plane, the algorithm works for single contacts in other directions as well. The case of multiple contacts is different, and assessment of the capability to restrict contact forces in such a scenario is left as future work.

6.6 Phantom study

6.6.1 Experimental setup and experimental procedure

The proposed compliant motion controller is further validated in an aortic phantom (Fig. 6.13). The phantom study was executed using a real-time LSTM controller, implemented in Python with the PyTorch library and GPU acceleration. Inter-component communication was facilitated by the Robot Operating System (ROS) and Orocos [284]. Specifically, the LSTM was integrated as a ROS node, which received the real-time position of the catheter tip from the EM sensor node. The output of the LSTM node was subsequently directed to control the pressure valve. The entire system operates at a frequency of 40 Hz.

The phantom features the ascending aorta, the aortic arch, and the descending aorta. The phantom is made of wood and is manufactured with laser cutting after projecting a real aortic model onto a two-dimensional plane. The phantom was connected on top of the F/T sensor (see inset of Fig. 6.13) so as to measure the force applied by the catheter on the vessel wall. A catheter, with one steerable segment and several passive segments, called 3Flex [285] is employed in this study. A camera (Allied Vision Technologies GmbH, Germany) is installed directly above the phantom to record the catheter motion. Two EM sensors are embedded at the tip and at the base of the first catheter segment. The 3Flex catheter is inserted into the descending aorta, advances through the aortic arch, and is carefully maneuvered until it reaches the aortic root. Afterwards, the catheter is retracted. The above procedure is carried out both with and without employing the compliant motion controller.

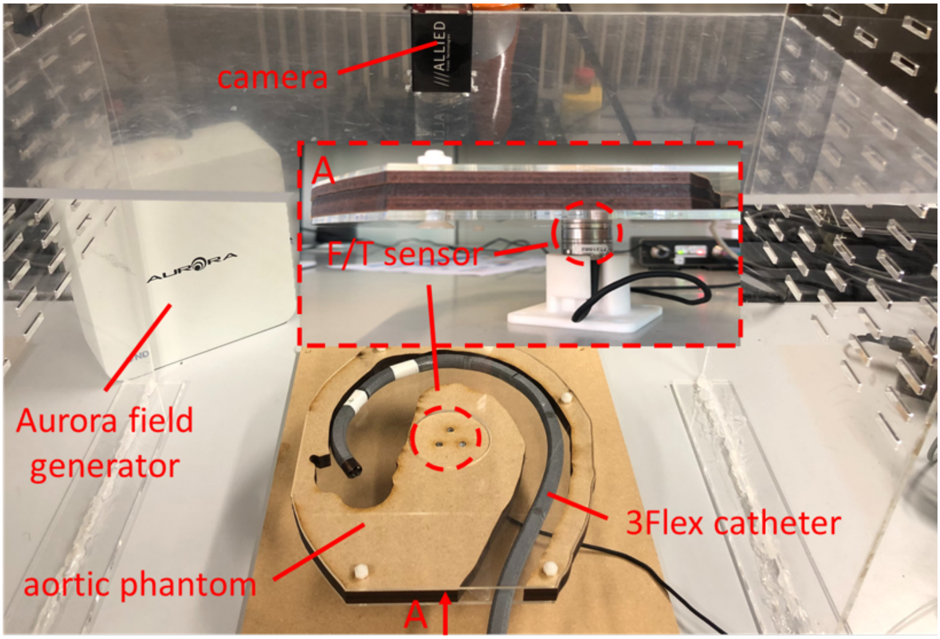


Figure 6.13: The experimental setup for the phantom study features a 3Flex robotic catheter equipped with an EM sensor at its tip, a laser-cut aortic phantom, and a F/T sensor mounted beneath the phantom to capture interaction forces. Additionally, an overhead camera records the experimental procedure, and an EM tracking system that includes the Aurora field generator is also integrated.

6.6.2 Results and discussion of the phantom study

During insertion, the average force and maximum force achieved by the compliant motion control are 0.14 N and 0.43 N, respectively, which are 65.8% and 70.7% lower than without the compliant motion control (Fig. 6.14). The catheter can bend to reduce the interaction force and move smoothly until it reaches the aortic root. On the contrary, when the proposed controller is off, the catheter is stuck in the aortic arch with its tip directly puncturing the vessel wall. During retraction, the catheter can even not be retracted without the compliant motion control, causing a maximum force of 0.73 N. When the compliant motion control is running, the catheter can be retracted without too much effort, while maintaining the interaction force below 0.44 N. This is 40.0% lower than without the controller running. The results also show that the method could reduce the user's mental load as the user does not need to focus on the force control, which allows the user to pay more attention to high-level tasks such as decision making. In addition, a reduction in the procedure duration is also observed with the method.

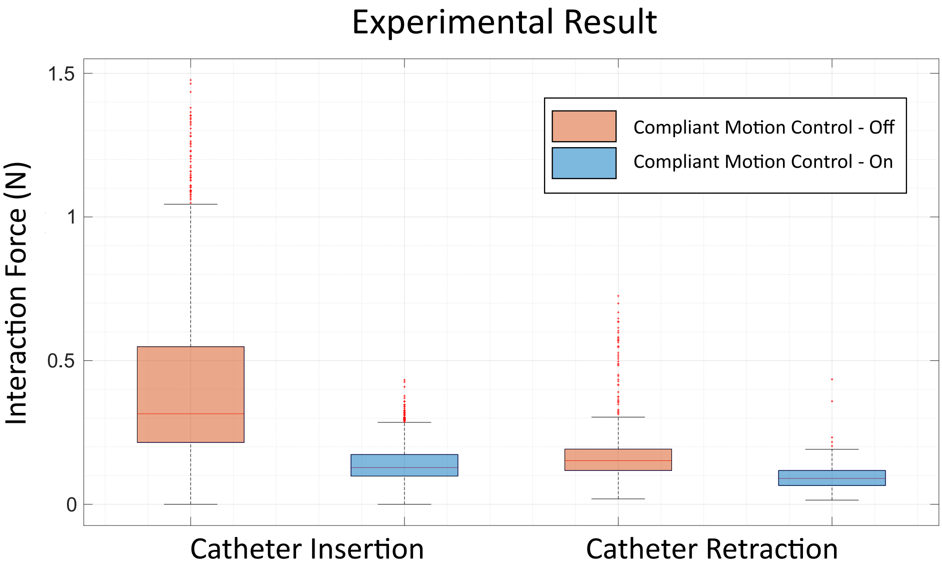


Figure 6.14: Box plot illustrating the interaction forces experienced during the insertion and retraction phases in the phantom study.

6.7 Conclusion on the DL-based compliant motion controller

This chapter introduced a deep-learning-based compliant motion controller for a robotic catheter that is aligned with sub-objective 5. Such a controller can actively bend the catheter tip away from the environment so that excessive interaction forces can be automatically avoided. This feature could contribute to safer interventions. The proposed method only requires one sensor that measures the tip position. The approach is simple and does not require complex sensors that measure distributed contacts or that measure the input torque applied by the actuators. The proposed compliant motion controller features a tip position controller based on a control-LSTM. This work shows how the control-LSTM was able to compensate for rate-dependent, saturated, and asymmetric hysteresis present in the catheter actuation. Trajectory following experiments revealed that the control-LSTM achieved sub-degree precision. The controller surpassed the analytic IDRDPI model by a minimum of 75% across all four test paths. Based on this tip position controller, a compliant motion controller was proposed. Experiments both in a simplified setup as well in an aortic phantom showed how catheter contact forces could be reduced by approximately 70%. Although showcased on a pneumatically-driven catheter, this method is believed to be transferable to other actuation methods such as cable-driven with moderate extra efforts. This work shows that it is not always necessary to elaborate complex mechanical models to steer robotic catheters.

Chapter 7

Valorization plan of the research results

7.1 Summary of the research results

7.1.1 Introduction

The current research environment is increasingly focused not only on uncovering new technologies but also on transforming this knowledge into real advantages for society. This process, commonly known as “valorization”, ensures that the results of research are not limited to academic publications, but are used to create products, services, or policies, thus maximizing the effect of research. Recognizing its importance, this chapter is dedicated to outlining the valorization plan for the research results obtained in this thesis.

7.1.2 Summary of the research results

Enhanced visual feedback and teleoperation: In current practice of endovascular interventions, catheters are inserted manually by clinicians under fluoroscopy guidance. Unfortunately, this approach introduces radiation to clinicians. Tissue damage may also occur, especially in the depth direction, since fluoroscopy is a 2D imaging modality thus lacks depth perception. This work investigates and compares different interactive modalities, namely teleoperation via a gamepad versus gesture-based steering via a Microsoft Hololens. The differences in visualization methods, namely via Hololens or via a 2D monitor, were also explored. A user study was conducted in which users teleoperated the catheters with different interactive modalities to follow a predefined path in an aortic phantom. Experimental results show that the combination of gamepad and Hololens achieved the most popularity among the users and the smallest trajectory tracking error. The superiority of the gamepad is the passive haptic feedback offered by the device. Whereas in the Hololens the user needs to visually confirm the input motion that was commanded. When using a gamepad, the user can rely on his/her proprioception to understand the pose and relative displacement of the input device. The HoloLens, however, provide 3D visualization and interactions with the holographic images compared to a 2D monitor.

Valorization potential: This chapter discusses methods that can enhance existing visual feedback in robotic catheter systems. Two primary improvements are revealed: first, a radiative imaging technique is replaced with a non-radiative method, and second, utilizing 3D visualization to better represent the catheter shape and the vessel environment, thereby improving depth perception of physicians. Currently, most robotic catheter systems on the market use teleoperation control methods. While this has changed the traditional model, where physicians had to steer the catheters directly beside the patient, these systems still rely on fluoroscopy for visual feedback. This still presents issues such as exposure to radiation and the limitations of 2D grayscale imaging. Therefore, the 3D visualization technology developed in

Chapter 3, utilizing AR-HMD, can be readily integrated with existing endovascular robotic systems on the market, such as Corindus (Siemens Healthineers, Germany). This integration would enable clinicians to use AR-HMD during interventions for a 3D view, as opposed to relying on traditional 2D monitors

Furthermore, Chapter 3 explores different control input devices. In current endovascular robotic systems, physicians typically control the catheter using joysticks and buttons fixed on a panel. However, display methods such as Hololens offer the possibility for users to interact 360 degrees with holographic images. Therefore, employing a portable gamepad as the control input device could further enhance the advantages of 3D visualization of HoloLens. Utilizing gamepad-based control can also be integrated with existing robotic systems on the market.

FBG-based shape sensing: A DL method for shape sensing of continuum robots based on multi-core FBG fiber is introduced. The proposed method, based on an ANN, differs from traditional approaches, where accurate shape reconstruction requires tedious characterization of many characteristic parameters. A further limitation of traditional approaches is that they require either multiple fibers whose location relative to the centerline must be precisely known (calibrated) or a single multicore fiber whose position typically coincides with the neutral line. The proposed method addresses this limitation and thus allows shape sensing based on a single multicore fiber placed off center. This helps in miniaturizing and leaves the central channel available for other purposes. The proposed approach was compared with a recent state-of-the-art model-based shape sensing approach. The superiority of the proposed approach was demonstrated. To address concerns regarding the “black box” nature of ANN, ablation study was conducted. This study offered deeper insights into both the structure and behavior of the ANN.

Valorization potential: The advantages of fiber-optic sensing technologies are many, offering benefits such as immunity to electromagnetic interference, high data sampling rates, and seamless system integration [65]. In recent years, interest in Fiber-Optic Sensing has surged in both the academic and commercial sectors. A case in point is the Fiber Optic RealShape (FORS) system of Phillips (Amsterdam, the Netherlands), developed to enhance instrument visualization while reducing the reliance on fluoroscopic imaging.

Similarly, the FBG technology adopted in this chapter can achieve comparable functionalities. FBG can be used for 3D shape reconstruction of catheters. One innovation presented in Chapter 4 is to overcome the conventional limitation that requires FBG sensors to be placed within the central channel of a catheter for accurate shape reconstruction. The findings indicate that even when FBGs are positioned on the outer surface of a catheter, precise shape reconstruction can still be achieved while maintaining high accuracy in both free and constrained spaces.

Currently, most cardiovascular robotic catheter systems still lack integrated shape-sensing capabilities and rely extensively on fluoroscopy, presenting an opportunity for collaborative advancement. Beyond vascular intervention robots, this technology holds promise for other types of robotic systems, such as bronchoscopy robots. To the best of the author's knowledge, there are few endoluminal robotic systems with integrated shape-sensing technology. One example is Ion system (Intuitve Surgical Inc., CA, USA). The Ion system offers a comprehensive view of the catheter in the airway through real-time 3D visualization, enabled by shape-sensing technology, while also providing 2D direct visualization via a vision probe at the catheter's tip. Therefore, collaboration with other bronchoscopy robot products and the integration of shape-sensing technology is also a potential valorization. Note that the FBG-based shape sensing technology is also compatible with passive or non-motorized instruments, and this integration is easier compared to the integration with a robotic system, especially in passive instruments that has an available channel for FBG. This integration facilitates the real-time visualization of the pose and motion of the instrument during endoluminal interventions.

Deep-learning-based hysteresis modeling and compensation: This method employs DL techniques to model and copensate for the hysteresis that appears in the catheters. For DL-based hysteresis modeling, a significant advantage of this method is its ability to predict the catheter's bending angle purely based on the input control commands, on condition that the catheter moves in free space. This method eliminates the need to install sensors at the tip of the catheter to determine the tip position. To implement this method, data must be collected from the catheter system, followed by training an ANN using these data on a GPU. This approach demonstrates the ability to forecast the tip bending angle with precision below one degree, outperforming traditional analytical hysteresis models. Additionally, this method was integrated with a Forward-Looking OCT fiber, illustrating the potential to reconstruct the environment ahead of the catheter using this combined approach. It is noteworthy that this method has been demonstrated to be capable of modeling hysteresis in catheter bending, but it can also be adapted to model the hysteresis in catheter rotation and translation. If the ANN is trained by swapping the input and output, it can also be adapted to compensate for hysteresis, thereby functioning effectively as an open-loop controller.

Valorization potential: In robot-assisted endovascular interventions, physicians primarily depend on visual feedback to compensate for hysteresis, which affects the intuitiveness of catheter steering and increases the cognitive burden of physicians. The hysteresis modeling and compensation methods presented in Chapters 5 and Chapter 6 offer solutions to these challenges. Moreover, these approaches are also adaptable to other endoluminal robotic systems, significantly enhancing their control precision. The method is ready to be implemented in current endoluminal robotic systems, such as Corindus, to compensate for any residual hysteresis. Although hysteresis is minimized in the initial design and manufacturing phases of catheters, it could still present in these systems. When an endoluminal robot is used for the first time, its input-output behavior is recorded and utilized to train two Artificial Neural Networks (ANNs). One ANN models the forward kinematics, mapping from control commands to catheter tip responses, while the other models the inverse kinematics, mapping from the desired catheter tip responses to corresponding control commands. This method effectively compensates for hysteresis in endoluminal robotic systems, thereby making teleoperation more intuitive and responsive.

Moreover, hysteresis is a pervasive issue that goes beyond continuum robots and is widely encountered in mechanical systems, sensors, actuators, magnetic materials, and shape-memory alloys [286]. Therefore, the methods and know-how presented in this chapter have broader applicability and can be used to address hysteresis in various domains.

Deep-learning-based safe and autonomous catheter control: In cardiovascular interventions, when using steerable catheters, it is necessary to take great care not to apply excessively large forces on the vessel walls, as this could dislodge calcifications, induce scars, or even cause perforation. To address this challenge, a novel compliant motion control algorithm was developed that relies solely on position sensing of the catheter tip and knowledge of catheter behavior, while a force sensor at the catheter tip is not needed. The proposed algorithm features a data-driven tip position controller. Experiments in both a simplified setup and an aortic phantom showed that the proposed approach allows the interaction forces with the environment to be reduced by around 70%. This research demonstrated the potential of DL to effectively manage the intricate nonlinear behavior of catheters, allowing for the smooth navigation of continuum robots in confined spaces such as the patient's blood vessels.

Valorization potential: Chapter 6 introduces a semi-autonomous control approach for catheters. Current commercial robotic catheter systems operate at Level 0 autonomy [25]. However, the trend is shifting towards incorporating higher levels of autonomy, which can significantly reduce the burden on physicians. Lower autonomy levels (1-2) assist in controlling and automating partial tasks, while levels above 3 enable complete automation of interventions, with physicians primarily in a supervisory role.

The method proposed in Chapter 6, subject to further optimization and thorough validation, has the potential to be licensed and integrated with existing endovascular robotic systems in the market, such as Corindus and R-one. Given that this control strategy is based on DL, a careful review of relevant policies and regulations is crucial prior to its commercial use. The envisioned clinical workflow is as follows: 1)For each robotic catheter used for the first time, data reflecting its input and output are collected using the strategy outlined in Chapter 6; 2)This data is employed to train an ANN; 3) The ANN will then serve as a controller for catheter steering, effectively functioning in both free and constrained spaces; 4)Physicians have the full freedom to switch between autonomous control and teleoperation with assistance as needed.

7.1.3 Data generated from the conducted research

The Ph.D. project aims to investigate catheter control based on DL. The data include raw experimental data, processed data, code, LaTeX files, and videos.

1. Software source codes are in Python, MATLAB, and C++.

2. The most important software/libraries that are used and/or adapted are MATLAB, ROS, Pytorch. All of them have either a permissive license or are open source.

3. All generated datasets were used mainly for scientific publications.

4. These datasets consist of: 1) data from simulations (for hysteresis), stored in CSV; (2) sensor measurements (pressure, catheter position, catheter shape, OCT data stored in CSV); (3) recorded videos of catheter motions or explanatory videos submitted as attachments to conferences or journals (stored in mp4).

5. Datasets and code are stored internally on local PC, Box/Onedrive or Gitlab/Github servers. The data are organized according to a directory structure that is based on the type of experiment. Access to data is granted on a need-to-know basis, and all access is logged and monitored.

6. There are no legal or ethical issues in the generated dataset. Data collected in user studies are anonymized to protect participant confidentiality.

7.2 Market survey

The market estimate and trends for catheters are summarized in the following Table 7.1 from [287]. It should be noted that the term “catheter” here is not limited to just vascular catheters, but also includes others such as urological catheters.

Table 7.1: Market estimate and trends for catheters.

Study	Value (year)/billion $	Estimate (year)/billion $	CAGR
Allied Market Research [288]	10.7 (2017)	20.0 (2025)	8.13%
Grand View Research [289]	37.3 (2018)	77.7 (2026)	9.61%
Market Study Report [290]	-	56.5 (2025)	-
Markets and Markets [291]	15.9 (2020)	24.2 (2025)	8.76%
Global Market Insights [292]	37.8 (2018)	79.0 (2025)	11.1%

Cardiovascular catheters are dominant in the catheter market, and are more prevalent than other types such as neurovascular and urological catheters. In 2019, the global market value for cardiovascular catheters was 10.9 billion USD. This value is projected to more than double, reaching 22.5 billion USD by 2026, highlighting a substantial market opportunity.

On a regional basis, North America is at the forefront in terms of revenue. Specifically, the US catheter market was valued at 10.8 billion USD in 2019, with forecasts suggesting a rise to 20 billion USD by 2026, as indicated by [289]. When considering these numbers, cardiovascular catheters represented a substantial portion, accounting for 3.17 billion USD in 2019, and this is expected to increase to 5.8 billion USD by 2026. Europe, while trailing the US, is making significant strides too. According to [291], the size of the European catheter market is on track to hit 6.6 billion USD by 2025. While there is notable variation in the market size estimates presented in Table 7.1, there is a unanimous consensus regarding the Compound Annual Growth Rate (CAGR). All sources project a CAGR of around 10%, which indicates a promising outlook for potential investors.

7.3 A valorization plan for sharing a public dataset

Introduction: The dataset, originating from the PhD research, describes the input-output behavior (pressure - tip bending angle or pressure - tip displacement) of a pneumatically actuated catheter system, which includes hysteresis. Hysteresis, observed across various disciplines, continues to be a focal point in contemporary research. Unraveling its complexities is essential to deepen insights in various fields.

Target users:

- Primary users: researchers within the medical or surgical robotics community.
- Extended users: researchers from other disciplines in which hysteresis is often encountered and needs to be addressed, such as magnetic materials, elastic materials, electrical circuits, thermodynamics, biology, and economics.

Value proposition: This dataset serves a manifold purpose:

- It provides a platform to test new hysteresis modeling algorithms.
- It accelerates research and innovations in the domain by providing readily available data.
- It eliminates the need for individual researchers to set up their own experimental systems, or preventing them from replicating established solutions.

Access, distribution and license:

- Access mode: the dataset will be open-access and made freely available. Additionally, sensitive portions might be restricted, shared on a need-to-know basis after consulting with PhD supervisors.
- Distribution responsibility: the main researchers and contributors oversee and distribute the dataset.
- Licensing: the data is under licenses such as Creative Commons, detailing its use, sharing, and potential commercial exploitation.

Infrastructure, maintenance, and enhancement:

- Platform: hosting is possibly on university databases.
- Maintenance: the owners guarantee that the dataset is maintained on a regular basis.
- Community contribution: researchers in the community are encouraged to contribute, making it a collaborative resource.
- Documentation & metadata: comprehensive metadata and documentation are provided to ensure correct usage and understanding.

Conclusion: by incorporating varied modes of valorization, from community involvement to technical structure, this plan ensures the broad utility of the dataset and maintains research integrity. The underlying priority remains to advance research while respecting data privacy and ethics.

7.4 Valorization plan for a DL-based hysteresis modeling and compensation algorithm

7.4.1 Description of the algorithm

The hysteresis modeling and compensation algorithm leverages DL techniques, primarily developed in Python and using the PyTorch library. Utilizing GPU acceleration, the ANN undergoes extensive training to ensure robustness and efficiency. This algorithm has a dual function:

- Modeling of hysteresis: It can accurately model the complex rate-dependent, asymmetric and saturated hysteresis in various systems.

- Compensation of hysteresis: Beyond mere modeling, the algorithm is adept at compensating for hysteresis and deriving the inverse kinematics of the robotic system and thus allowing precise control of the system.

7.4.2 Possible failure modes of the algorithm

- Hardware and computational limitations: Challenges such as GPU overload with delayed or inaccurate feedback, and computational limitations can significantly impede the real-time processing and effectiveness of DL models, especially in critical applications such as surgery.
- Data issues and model robustness: Issues like poor data quality, model overfitting, alongside robustness and generalization problems, can limit the practical application of models, hindering their performance in real-world scenarios.
- Over-reliance on the DL-based method by clinicians: In medical contexts, there is a risk of clinicians becoming overly reliant on the model's outputs. This can lead to problems when the model encounters untrained scenarios or produces incorrect outputs, underscoring the need for clinician awareness of the model's limitations.

7.4.3 Protection strategies

Given the proprietary nature of the algorithm, protecting it as a trade secret is a potential route. A software can be developed based on the algorithm without revealing its intricacies:

- Develop application software: instead of sharing the algorithm directly, software applications that leverage the algorithm could be developed. The software provides the benefits of the algorithm without revealing its core logic. Instead of licensing the algorithm, the software can be licensed, making sure that the licensing agreement prohibits reverse engineering or any other attempt to discover the underlying algorithm.
- Layered architecture: design the software in layers where the algorithm operates at the core, isolated from user-facing interfaces. This helps to protect the intricacies of the algorithm[293].
- Obfuscation: in software, obfuscation refers to the process of generating source or machine code that is challenging for both humans and computers to decipher [294, 295]. Therefore, it is considered to use code obfuscation techniques to further disguise the algorithm within the software, making it difficult to reverse engineer [294].

7.4.4 Strategy for algorithm protection and development phase

The plan for protecting the algorithm is described as follows:

At the beginning, the main emphasis will be on further improving the DL algorithm while ensuring that its design, features, and related data remain confidential. In addition to this progress, all modifications, advances, and essential discoveries will be carefully documented to create a precise timeline and basis. These data will only be accessible to essential research team members, and any external collaborators or consultants will be required to sign a non-disclosure agreement. At the same time, the transformation of the algorithm into software will start, taking precautions to make sure that the main algorithmic code is kept hidden and protected against reverse engineering. Strategies such as code obfuscation may be used for this purpose.

Once the software/algorithm reaches a mature and stable version, then copyrighting the software may be the next step, even though this will not protect the algorithmic concept itself. However, it will secure the particular expression and code of the software, providing legal defense against unapproved copying or circulation.

7.4.5 Plan of licensing scheme

This plan aims to take advantage of the valorization potential of the DL algorithm for increased precision in robotic catheter control, while carefully considering the complex areas of licensing and regulatory compliance.

After securing protection for the DL algorithm, either through copyrighting or keeping as know-how, the next critical step involves introducing it to the market. Licensing is a viable option that could be used to spread the algorithm's abilities without giving up any proprietary rights.

The process of obtaining a license starts with recognizing potential collaborators, particularly those companies that develop endovascular robotic system. These include Cordindus endovascular robot of Siemens Healthineers and R-ONE robotic catheter of Robocath Inc., but can also be companies producing other endoluminal robotic systems such as Auris Health, Inc. and Intuitive Surgical. An exclusive license agreement [296] with a leading company in this field (see Table 1.1) can be investigated, allowing them the rights to employ the algorithm and thereby improving their competitive position in the market. The right to exclusively license a product or service may be limited to a certain period of time or contingent on the achievement of certain objectives. After this, the opportunity to license the product or service to other interested parties can be granted.

A carefully crafted licensing agreement is essential. This agreement should outline the extent of utilization, the right to grant sub-licenses, integration assistance, performance criteria, duties concerning updates or changes, and financial matters.

It is good to include legal professionals who specialize in intellectual property and technology licensing from the beginning. Regular meetings between technical and licensing teams will be held to ensure seamless technology transfer and address any potential issues that may arise during the application of the transferred technology, thus protecting the interests of both parties.

Given the use of AI in medical contexts, it is essential to pay close attention to regulatory demands [297]. Strict compliance with the Medical Device Regulation (MDR) in Europe is mandatory, which means that device manufacturing partners must conduct a thorough validation and, possibly, clinical trials to demonstrate the safety and effectiveness of the licensed method. It is also essential to adhere to the GDPR if the algorithm needs to use patient information to train the ANN.

Chapter 8

Conclusion and discussion

This chapter begins with a detailed summary of the research outcomes achieved. It then proceeds to evaluate the fulfillment of each predefined sub-objective, assessing the methods used, discussing their limitations, and providing reflections on these aspects. Section 8.3 outlines directions for future work, and the wider impact of our findings within the field of surgical continuum robots. The last section offers some brief contemplations inspired by the research.

8.1 General thesis summary

The escalating prevalence of CVDs worldwide represents a pressing health concern. CVDs, including conditions such as coronary artery disease, heart failure, and arrhythmias, often require complex diagnostic and therapeutic procedures. As the leading cause of mortality worldwide, CVD underscores the urgent demand for effective treatment supported by advanced interventional instruments. Endovascular intervention is at the forefront of this medical evolution, with the aim of improving patient outcomes, reducing recovery time, and minimizing potential complications.

The current landscape of cardiology interventional devices is characterized by a variety of technological innovations and methodologies. However, challenges still exist. The primary challenges lie in control and visualization. For decades, catheter procedures have been visualized using 2D grayscale fluoroscopy images. Although fluoroscopy offers real-time visualization, it subjects patients and clinicians to ionizing radiation, which can pose health risks. Furthermore, these 2D images do not capture the intricate 3D architecture of vascular structures, making navigation a formidable task.

Modern robotic catheterization systems address part of the drawbacks of current procedures. Predominantly, these systems are teleoperated, translating the clinician's hand movements into catheter movements within the vasculature of patients. Despite technological advances, these systems come with inherent limitations. Many physicians find that the learning process is difficult and requires a significant amount of time to become proficient. Issues such as hysteresis compromise precision. Additionally, the lack of haptic feedback further prevents clinicians from having an intuitive manipulation. Furthermore, all endovascular robotic systems on the market currently lack autonomy. This not only increases the physical demands on clinicians, but also makes the entire process highly dependent on the skill and experience of the clinicians.

Addressing these challenges, the research presented in this dissertation embarked on a comprehensive journey to further optimize robotic catheterization. The first contribution (presented in Chapter 2) involved performing a survey of the applications of ML, including modern DL techniques and traditional ML techniques, in the scope of FSIR.

The second contribution (presented in Chapter 3) investigates and compares different teleoperated control strategies, namely teleoperation via a gamepad versus gesture-based steering via a Microsoft Hololens. Differences in visualization methods, namely via Hololens or via a 2D monitor, were also explored. A user study was conducted in which users teleoperated the catheters with different interactive modalities to follow a predefined path in *in-vitro* environment. Experimental results show that the combination of gamepad and Hololens achieved the most popularity among the users and the smallest trajectory tracking error.

During vascular interventions, knowing the catheter status would be highly beneficial. For this reason, advanced sensing techniques were investigated in Chapter 4, primarily aimed at monitoring catheter status and reducing radiation exposure from fluoroscopy. The FBG sensor can reconstruct the shape of continuum surgical robots, although the accuracy of the FBG-based reconstruction is affected by several factors. First, the FBG-based reconstruction heavily relies on parameters provided by the manufacturers, which may vary due to the assembly processes. Furthermore, conventional deployment requires that the FBG sensors be placed in the central channel of the robot. However, certain robots lack this channel or use it for different functionalities. To avoid these issues, a DL approach that correlates wavelength shifts with the curvature and bending plane angles at each grating is introduced. This method offers more accurate shape reconstruction than traditional strategies.

One of the challenges in catheterization is hysteresis. Typically, clinicians compensate for hardware hysteresis purely by relying on their visual feedback. The research presented in Chapter 5 introduced innovative recurrent neural networks to minimize this phenomenon, ensuring predictable catheter movements. This method offers a versatile solution applicable to a variety of endoluminal robotic systems and even other disciplines, and setting the stage for increased level of autonomy.

Using the power of DL, a compliant motion control algorithm was developed in Chapter 6. This elevated the autonomy of the catheter to level 2 [25] with a demonstration of navigating in a 2D phantom. A significant highlight was the capability of the controller to autonomously reduce interaction forces between the catheter tip and the vessel, ensuring safer interventions.

This dissertation focuses primarily on the use of DL and AR techniques to elevate the efficacy of robotic catheterization procedures. Through meticulous development, the work has presented a synthesis of the cutting-edge technologies described above to address the challenges in current catheterization procedure. The dissertation presents bench-top experiments across different sub-objectives to ensure robust validation. Significantly, at least two of these experimental evaluations were conducted in an aortic phantom, demonstrating the applicability of the proposed methods. Although the current scope of the dissertation remains within *in-vitro*, it is planned to expand this research into *ex-vivo* or *in-vivo* environments in the subsequent phases, paving the way for potential clinical applications.

8.2 Analysis of the achievement of research objectives and limitations

Sub-objective 1: *Identification of the latest trends of ML techniques in FSIR.*

Evaluation of the results achieved: In alignment with the first sub-objective, a comprehensive survey was conducted on the application of ML techniques in FSIRs. Following the trend that the application of ML grows in various fields, its role in FSIR has also become increasingly pivotal. Specifically, ML has found applications in various tasks in FSIR, ranging from perception and modeling to advanced navigation and control. The review conducted describes a landscape of increasing adoption of diverse ML algorithms across different applications. Such insights not only indicate how ML advances might propel adoption in intraluminal procedures but also shed light on the potential functionalities that ML can offer. This review initially offers valuable insights into algorithm selection for the work presented in this thesis. In addition, it could serve as a comprehensive guide for practitioners and researchers in the field of FISR, offering recommendations for optimal algorithmic choices for their specific cases. This guidance facilitates the implementation of ML approaches in the field of FSIR, positioning them as effective alternatives to traditional analytical models.

Limitations: Although the survey provides a comprehensive overview of the application of ML in FSIR, it is not without limitations. First, due to the absence of clinician involvement, this chapter predominantly adopts an engineering perspective. The distinct clinical needs of various interventions, which are crucial for the implementation and adoption of FSIR, are not comprehensively presented. Moreover, visual elements such as images or illustrations of robots and experimental setups from the referenced papers are omitted. This might hinder readers from grasping the design of the robots or the specific experimental setup, necessitating a deeper dive into the original sources. Moreover, a clearer guidance detailing the appropriateness of specific algorithms for certain tasks would have further elevated the utility of the review. Lastly, the rapidly evolving ML field constantly witnesses the emergence of transformative technologies. This review does not delve deeply into these novel ML technologies that have the potential to revolutionize the FSIR landscape. Please note that this review is limited to works employing traditional ML, DL, or RL methods. It does not include those that are data-driven but utilize "gray-box" approaches such as [139]. Such methods, which potentially offer a high degree of interpretability, would be a worthwhile subject for a separate dedicated review.

Sub-objective 2: *Investigation of the most effective interactive modality for teleoperated endovascular intervention.*

Evaluation of the results achieved: A user study was conducted to compare different interactive modes in human-in-the-loop teleoperated endovascular intervention. Three interactive modes were implemented with different control input devices and visualization devices. Implementing validation through *in-vitro* vessel navigation presents significant challenges. These challenges stem from the limited availability of robust steerable catheter systems, the recent presence of high-quality AR-HMDs, and the complex requirements for integrating these components. Nevertheless, an AR user interface was developed, which facilitates not only catheter control, but also visual feedback. Another important preparation was the integration of an endovascular catheterization system, which included a robotic catheter, a HoloLens, and a gamepad for experimental validation. The findings of the user study suggest a clear preference for the use of the gamepad as a control input device and HoloLens as a visualization tool. The gamepad emerged as an intuitive, responsive, and portable control device, while HoloLens offered the possibility to allow users to navigate and change perspectives in physical space to interact with holographic images.

Limitations: One aspect to note is that the study focused solely on the intrinsic differences between 2D and 3D visualization, as the guidance cues were kept consistent across both methods. It is worth further investigating the differences in guidance cues. For example, our visualization strategy, which shows the catheter's tip pose with an arrow, might not be intuitive enough. Although this approach is easy to implement, it potentially leaves out crucial information such as the shape of the catheter during navigation. Using technologies such as FBG-based shape sensing can enable visualization of the complete shape of the catheter. This provides physicians with more information during catheter navigation and thereby enhancing the safety of interventional procedures. Additionally, the phantom employed in this study may not accurately reflect the vascular dynamics observed in the vessels. For a closer approximation to real-world conditions, it would be beneficial to conduct tests using phantoms filled with blood mimicking fluids, simulating the motion and characteristics of blood flow. Finally, the scope of the user study, in terms of the number of trials and the diversity of the participants, was also limited. Expanding the trials by involving a more extensive set of participants, especially clinicians, would offer richer insights. Moreover, advancing this study into *in-vivo* animal experiments could be a valuable extension, presenting a more holistic understanding of the capabilities and challenges of the system.

Sub-objective 3: *Determination of the 3D catheter shape using a non-radiative method.*

Evaluation of the results achieved: A DL approach for improved shape sensing of continuum robots utilizing a multi-core fiber FBG sensor was introduced. Traditionally, precise shape reconstruction necessitated extensive characterization of numerous parameters, and most methods demanded either multiple fibers with accurately known relative locations or a singular multi-core fiber aligned with the neutral line. The proposed method overcomes these restrictions, allowing for shape sensing with an off-center multi-core fiber. This opens up an opportunity for miniaturization and retains the central channel of the robot available for other utilities. Upon comparison with a recent state-of-the-art model-based shape-sensing technique, the superiority of the proposed method became evident. Additionally, to improve the interpretability of the ANN, multiple ablation studies were conducted. Note that the technologies developed in this chapter are not only compatible with robotic systems but can also be applied to passive or non-motorized instruments. This allows for real-time acquisition of the instrument's shape. Furthermore, when combined with the advancements presented in Chapter 3, the instrument's shape can be visualized in an Augmented Reality Head-Mounted Display (AR-HMD), enabling 3D visualization.

Limitations: Firstly, despite the advancements, its foundation on an ANN, commonly viewed as a "black box", could raise concerns of limited interpretability. While our ablation studies do improve the interpretability of the ANN, the use of Explainable AI models [298] could provide even greater transparency and effectively address this issue. Furthermore, the DL approach, though compared with conventional shape reconstruction techniques, has not yet been benchmarked against traditional ML methods. Such a comparison could offer more comprehensive insights into its advantages or potential areas of improvement. In addition, the datasets used for training were built on the constant curvature assumption derived from the catheter tip and base poses. In future work, alternative sensing modalities, such as shape reconstruction from silhouette, could be considered for training data generation. This would pave the way for the application of the proposed method to other flexible instruments that may not strictly adhere to the constant curvature assumption during training. Lastly, while demonstrating the superior accuracy of our method in shape reconstruction is an important achievement, its full impact can only be assessed when integrated into the control system. Evaluating how these advances in shape sensing could influence the accuracy of catheter control will be critical for understanding the comprehensive benefits of improved shape sensing.

Sub-objective 4: *Identification and compensation of non-linearities in catheters.*

Evaluation of the results achieved: In line with the sub-objective to model and compensate for the non-linearities in a catheter system, substantial progress was achieved. The research successfully demonstrated the potential of employing AI, specifically DL, to model hysteresis, a phenomenon that is inherently rate-dependent, asymmetric, and saturated. The intricacies of this hysteresis are sourced from various factors, including the actuation method, friction between different catheter components, the compliant nature of materials, and the non-linearity of pressure valves. The proposed DL method is able to model all these contributory elements in one shot. This model effectively predicted the bending angle of the catheters with sub-degree precision, outperforming traditional analytical models. This accomplishment was further enhanced by integrating the model with a Forward-

Looking OCT fiber, enabling real-time environmental reconstruction directly in front of the catheter tip. This method does not need dedicated sensors at the catheter tip for localization. Additionally, an open-loop controller that uses a control-LSTM has been designed to manage and compensate for the hysteresis in the catheter system. The control-LSTM is developed by training the LSTM with reversed input-output mappings. The proposed control-LSTM serves as a robust feedforward controller for catheter motion in free space, effectively eliminating the influence of severe hysteresis. The control-LSTM is precise and has a simple training procedure. Moreover, it is a direct inverse mode. In other words, the control-LSTM does not require a separate inversion step to be used in a controller, as would be required by most traditional analytical hysteresis models, such as [139].

Limitations: Despite these advances, certain limitations underscore the future direction of the research. The validation was solely based on a pneumatically driven system, leaving exploration in the realm of cable-driven systems as a future work. It would be valuable to investigate whether the hysteresis pattern observed in a cable-driven system differs substantially from that in a pneumatically driven system, and how well the suggested DL model can be adjusted and used for different patterns of hysteresis. In addition, the study only focused on a 1-DOF unidirectional and a bidirectional system. This prompts an exploration into a system with spatial bending capabilities (2-DOF). Theoretically, 2-DOF bending would make the issue of dead zones more complicated, which increases the efforts of modeling. Finally, the interpretability of the ANN could be further improved. This can be achieved either through an ablation study, or by incorporating explainable AI models to provide more transparent insights into its decision-making process [299].

Moreover, in endovascular interventions, hysteresis is present in each DOF of the catheter, including translation, rotation, and bending. However, the findings in Chapter 5 primarily address bending hysteresis. Translational hysteresis is often a result of complex vascular anatomy. For instance, during interventions, a physician may encounter a delay when changing the motion direction of catheter. Catheter only becomes responsive again after a certain distance of insertion or retraction. Rotation hysteresis refers to the non-linear relationship between the rotation at the catheter tip and its proximal end. Addressing translational hysteresis presents a greater challenge as it is influenced not only by the catheter's characteristics but also by the surrounding environment. The LSTM method, as applied in Chapter 5, should be further developed to effectively tackle hysteresis in both translation and rotation.

Sub-objective 5: *Elevation of the level of autonomy in robotic catheters, along with its demonstration in an in-vitro environment.*

Evaluation of the results achieved: In this study, a novel compliant motion control algorithm, which operates solely based on the position sensing of the catheter tip and the modeling of the catheter behavior, was introduced. Central to our method is the data-driven tip position controller that is trained based on the "control Long Short-Term Memory Network" (control-LSTM). To determine the efficacy of the control-LSTM, trajectory following experiments were carried out on four distinct trajectories. The control-LSTM exhibited superior positioning abilities, managing sub-degree precision while robust to rate-dependent hysteresis. Both simplified setups and aortic phantom experiments highlighted the effectiveness of the proposed method in reducing the interaction forces with the environment by 70%. This underscores the potential of leveraging DL techniques to avoid exhaustive modeling that was traditionally needed for accurately navigating continuum robots in intricate environments such as the patient's vasculature.

Limitations: Despite the advances, the study has its set of challenges. Validation was limited to a 2D phantom. Extending this to a 3D phantom filled with blood mimicking fluid to replicate blood flow dynamics would undoubtedly enhance the practical relevance of our results. Moreover, our compliant motion control was exclusively operational in a single bending plane. Given that catheters in actual clinical settings encounter forces from multiple directions, expanding our method to accommodate forces in 2DOF becomes imperative. The manual insertion method in our validation is another point for potential improvement. True autonomy in navigation would mean a combination of both autonomous insertion and bending, preferably integrated with a catheter driver. Building on the expertise garnered from the research presented in Chapter 3, integrating a catheter driver into validation does not pose challenges. In terms of sensing, the current feedback mechanism relies on an EM sensor, which offers limited control information. In contrast, the use of an FBG sensor could substantially expand the range of information available. This FBG can not only determine the tip position, but also provides a comprehensive understanding of the catheter's shape, along with the magnitude and location of forces applied to it. This information can be beneficial for evaluating the extent to which interaction forces have been reduced. Lastly, increasing the inference speed of the ANN could potentially enhance the bandwidth of the controller, thus offering more real-time control responses.

8.3 Future directions

8.3.1 DL-based catheter control under environmental contact

This thesis has showcased the promise of DL techniques for achieving precise control over the distal segment of flexible/continuum robots in free space [138, 239]. This approach holds significant utility in specific clinical scenarios, for example, within the heart. However, there are a few instances during cardiovascular interventions in which the catheter comes into contact with blood vessels and other soft tissues. In these cases, the performance of an ANN trained on data collected in free space scenarios would inevitably deteriorate.

Past efforts, including those of this author, to train different ANNs with obstacles present at varying locations have encountered challenges. Such attempts require a large amount of data, a tedious data collection procedure, and significant training time, yet they offer limited generalizability [300]. The potential of ANN to improve catheter control can be further explored. A notable shift from our previous strategy is that the structure of the ANN can be adjusted to include the obstacle's position as a continuous input. Data collection is carried out with obstacles placed at different locations. During the testing, the FBG sensor is used to first estimate the location of contact, following the methodology of [128]. The position of the predicted contact point and the desired trajectory are then fed into the ANN, allowing it to forecast the appropriate control commands. This methodology will first undergo verification in a 1-DOF setting, followed by an expansion to a more complex 2-DOF spatial bending. This strategy is expected to pave the way for accurate control of the varied instruments when they interact with their surroundings, eliminating the need for embedded sensors at the distal end.

8.3.2 Validation of the hysteresis compensation method on catheters with other types of actuation

Cable-driven catheter system: DL methodologies were used to model and compensate for hysteresis in pneumatically-driven catheter systems. The hysteresis phenomenon predominantly results from the intricate nonlinear behavior stemming from the compressibility of the air, friction between the balloon and the braid, and the non-linear nature of the soft materials used to build PAM.

Nevertheless, when applying this DL-based approach to cable-driven robotic systems, it presents a new set of difficulties. Cable-driven mechanisms present hysteresis rooted in distinct characteristics, such as cable stretch under tension and friction as cables navigate through guides or pulleys. The intricate interplay between cable tensioning, frictional forces, and potential backlash may require an evolved model that can capture these nuances. Furthermore, managing phenomena such as cable slack will demand enhanced adaptability from our models. Drawing inspiration from our previous successes with pneumatic actuation, our future endeavors will be geared towards tailoring and validating our DL models to address these unique challenges posed by cable-driven systems such as those in [301, 302]. The overarching goal remains consistent: ensuring precise and predictable catheter control by effectively compensating for hysteresis, regardless of the actuation method.

Magnetically-driven catheter system: Magnetically-driven catheter systems [303, 304] are gaining significant traction nowadays. These systems, whether in micro- or standard-size, leverage magnetic fields to induce motion, making them especially suited for intricate surgical procedures where non-contact actuation is beneficial.

Magnetically-driven catheters offer enhanced maneuverability, reduced mechanical friction, and potential for miniaturization, while mitigating wear-and-tear challenges present in mechanically-actuated systems. However, magnetic actuation brings its own set of complexities, the main of which being hysteresis [305].

The hysteresis observed in magnetic-driven systems often arises from the nonlinear relationship between the magnetic field strength and the induced motion. Factors such as varying distances between the magnet and the catheter, magnetic field interference, and magnetic material characteristics can introduce unpredictability in the catheter's response. This unpredictable behavior further underscores the need for sophisticated modeling techniques, such as the DL approaches that we have used previously. With our established foundation in compensating for hysteresis in pneumatic systems, it stands to reason that adapting and refining our methodologies for magnetic-driven catheters is a promising and logical next step, ensuring that the benefits of non-contact actuation are not overshadowed by control challenges.

8.3.3 Integration of technologies and developmental outcomes from this thesis

The developments presented across various chapters of this thesis hold significant potential for integration, paving the way for a more comprehensive system for endovascular interventions. The integrated system was envisioned to involve the use of a gamepad for remote catheter steering, complemented by visual feedback through an AR-HMD as implemented in Chapter 3. This system would offer real-time updates on the pose and motion of the catheter, as well as the surrounding environment, thereby improving the control and situational awareness of physicians during procedures.

In such a system, a robotic catheter is manipulated by a dedicated catheter driver, supported by a serial robotic arm to adjust insertion angles for optimal navigation. Following further development of the semi-autonomous control method presented in Chapter 6, the system could have varying levels of autonomy tailored to the complexity of different procedures. For simpler tasks, such as aortic navigation, the system could operate autonomously, requiring the physician to only monitor the procedure. More complex procedures, such as navigating the coronary arteries, would utilize teleoperation for greater safety and precision.

Before each robotic catheter is used, its input-output data are meticulously collected for the purpose of training an ANN as introduce in Chapter 5 and Chapter 6. This ANN is adept at compensating for hysteresis of a catheter and effectively establishing a robust kinematic model, thereby enhancing the catheter's functionality. Additionally, incorporating the FBG sensor within the catheter for real-time shape sensing, as implemented in Chapter 4, would allow the physician to visualize the catheter position and configuration in 3D through AR-HMD. Combined with pre-operative CT scans and intraoperative imaging such as IVUS or OCT, this would provide a detailed and constantly updated view of the vascular environment.

The clinical workflow envisioned for the integrated system, with a focus on CTO treatment, is outlined as follows.

- Utilizing ML techniques to analyze patient-specific data and previous cases, identifying the latest trends and optimal approaches for the specific intervention.
- Planning the intervention using 3D models of the patient's vasculature, constructed using pre-operative CT. The physician can view the patient's anatomy in an AR-HMD to gain a comprehensive 3D visualization.
- The assistant will collect the input-output behavior of the catheter to be used and utilize this data to train various ANNs for hysteresis compensation and kinematic modeling. The assistant will interact with a GUI, focusing on user-friendly operation without delving into technical details.
- Preparing the robotic catheter system, ensuring it is equipped with an interface for tele-operated control using the chosen effective interactive modality, and another option for (semi-) autonomous control.
- Vascular access is typically gained through the radial or femoral artery. For illustration, accessing via the femoral artery is used as an example. Catheters and guidewires are inserted into the femoral arteries.
- Initiating the procedure with semi-autonomous or fully autonomous modes for navigating the descending aorta, aortic arch, and ascending aorta. The physician supervises the entire procedure while wearing an AR-HMD, which displays both the 3D shape of the catheter reconstructed by FBG and the vessel shape reconstructed from the IVUS probe at the catheter tip.
- Once the catheter reaches the aortic root, delivery catheter is anchored. The aortic root serves as the entry point to the coronary arteries.
- Deploying a micro-catheter/guidewire from the delivery catheter and switching to teleoperation modes for navigating the coronary arteries. The physician continues to receive 3D visualization through AR-HMD, displaying the catheter status and the surrounding vessel environment.
- Navigate the microcatheter/guidewire through the coronary artery to reach and cross the CTO. The selection of techniques and tools varies based on the chosen approach (antegrade or retrograde) and the specific characteristics of the occlusion.
- Following the crossing of the CTO, retract the microcatheter and the delivery catheter.

- Conduct further imaging to confirm the restoration of adequate blood flow so as to evaluate the performance of treatment.

However, such integration faces considerable challenges, particularly in minimizing latency in the visualization pipeline for effective real-time feedback. Overcoming these challenges is important for the successful application of these advanced technologies in enhancing the safety and effectiveness of endovascular interventions. New hardware design is also anticipated to play a key role in realizing the above clinical workflow for treating CTO.

8.3.4 Towards in vivo experimental validation

The advances that were validated through the phantom study are noteworthy. However, the next logical progression in our research trajectory is to bridge the gap between phantoms and real-world scenarios through animal trials. While the aortic phantom has provided invaluable insights and a controlled environment, it does not replicate the dynamic intricacies of blood flow, tissue responses, or the complexity of live surgical procedures.

Given the context and requirements of our trials, the pig may be the best choice for our animal experiments. Their significant anatomical and physiological similarities to humans make them ideal [306]. The cardiovascular system of the pigs is particularly suitable to test our catheterization methods. The fact that young or small pig breeds are compatible with the same catheters used for humans suggests that our findings can be more easily applied to human procedures.

Embarking on these trials involves multiple preparatory steps. First, ethical clearance from a recognized committee will be applied. This process is stringent, ensuring the welfare of the animals. It would require us to meticulously detail the purpose of our research, the number of pigs involved, and the exact procedures. In addition, our commitment to minimizing pain and ensuring the well-being of animals must be demonstrated. Once approved, the selection of the pigs becomes important. Healthy pigs without conditions are essential. Their housing, nutrition, and post-operative care, both before and after procedures, must be of the highest standards to ensure their well-being and the validity of our findings.

The process will be designed to mimic human catheterization as closely as possible. Pigs will be placed under anesthesia in a suitable manner, and our catheterization methods will be used with careful, ongoing observation. This will provide immediate feedback on our methods and the animals' responses. After the procedure, a comprehensive evaluation is necessary. This will involve examining any tissue responses at the catheter insertion sites, and evaluating the reactions of the cardiovascular system.

In conclusion, *in-vivo* experiments could provide insight that will not only validate current methodologies, but also lay a strong foundation for future endeavors, progressing towards human clinical trials.

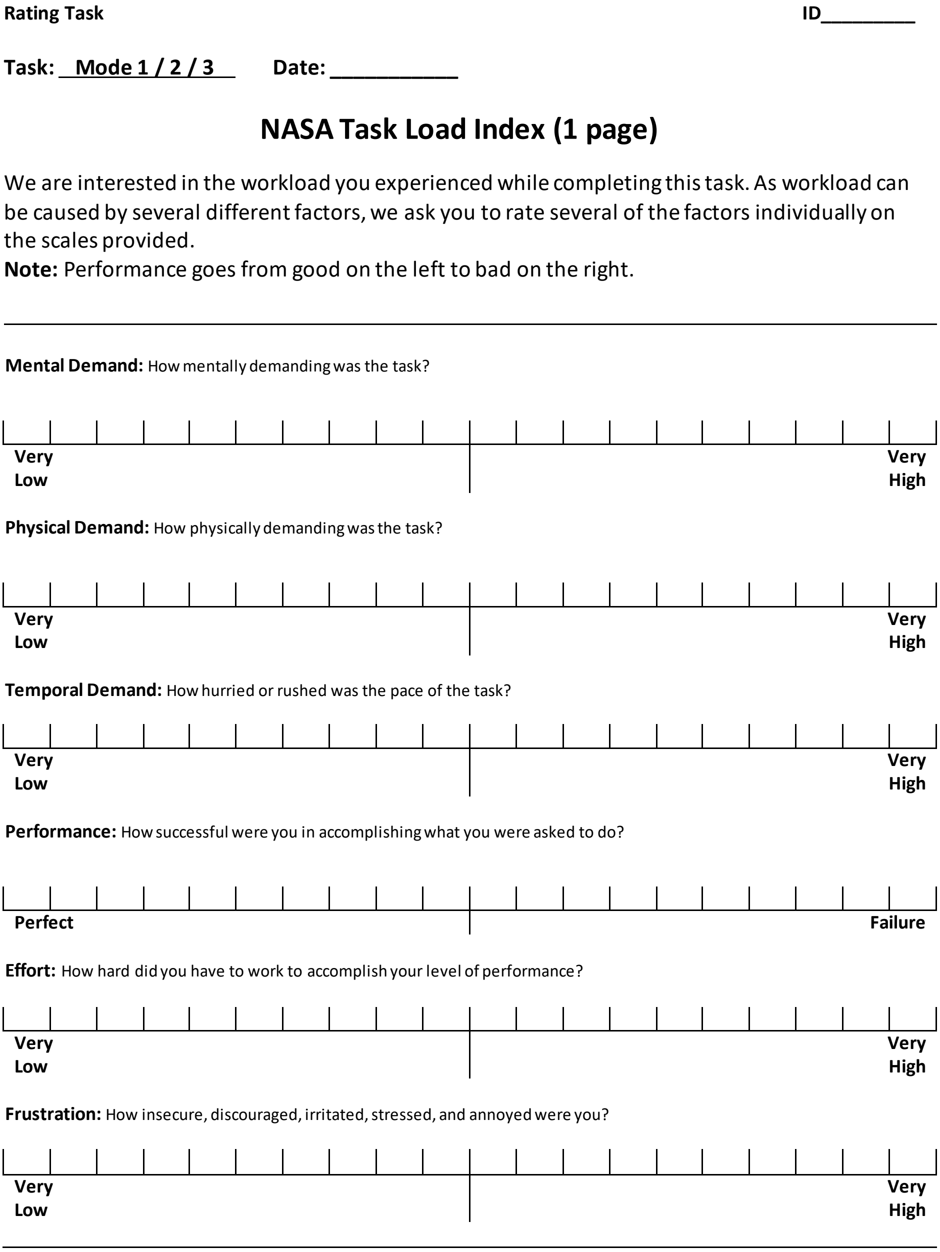
8.4 Concluding remarks

This thesis has demonstrated the far-reaching potential of using DL and augmented reality techniques to advance the control and sensing of robotic catheters. One critical lesson learned is the importance of remaining open to emerging technologies, even those with limitations and doubts. Taking DL for instance, while concerns over its interpretability remain, this should not deter us from employing its predictive and modeling capabilities. Technologies evolve, and advancements such as explainable AI are already emerging to address these limitations.

It is also crucial to learn that DL is not a one-size-fits-all solution. Overuse of DL is pointless. It is recommended to systematically compare DL approaches with model-based methods on the same subject matter. Following this comparison, the decision to employ DL should be carefully evaluated, considering key metrics such as performance and interpretability for each approach.

Appendix A

NASA Task Load Index



Rating Task **ID________**

Task: Mode 1 / 2 / 3 **Date: ___________**

NASA Task Load Index (1 page)

We are interested in the workload you experienced while completing this task. As workload can be caused by several different factors, we ask you to rate several of the factors individually on the scales provided.
Note: Performance goes from good on the left to bad on the right.

Mental Demand: How mentally demanding was the task?

Very Low **Very High**

Physical Demand: How physically demanding was the task?

Very Low **Very High**

Temporal Demand: How hurried or rushed was the pace of the task?

Very Low **Very High**

Performance: How successful were you in accomplishing what you were asked to do?

Perfect **Failure**

Effort: How hard did you have to work to accomplish your level of performance?

Very Low **Very High**

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low **Very High**

Figure A.1: NASA Task Load Index, a subjective workload assessment form designed to measure the perceived workload of individuals performing a task, adapted from [205].

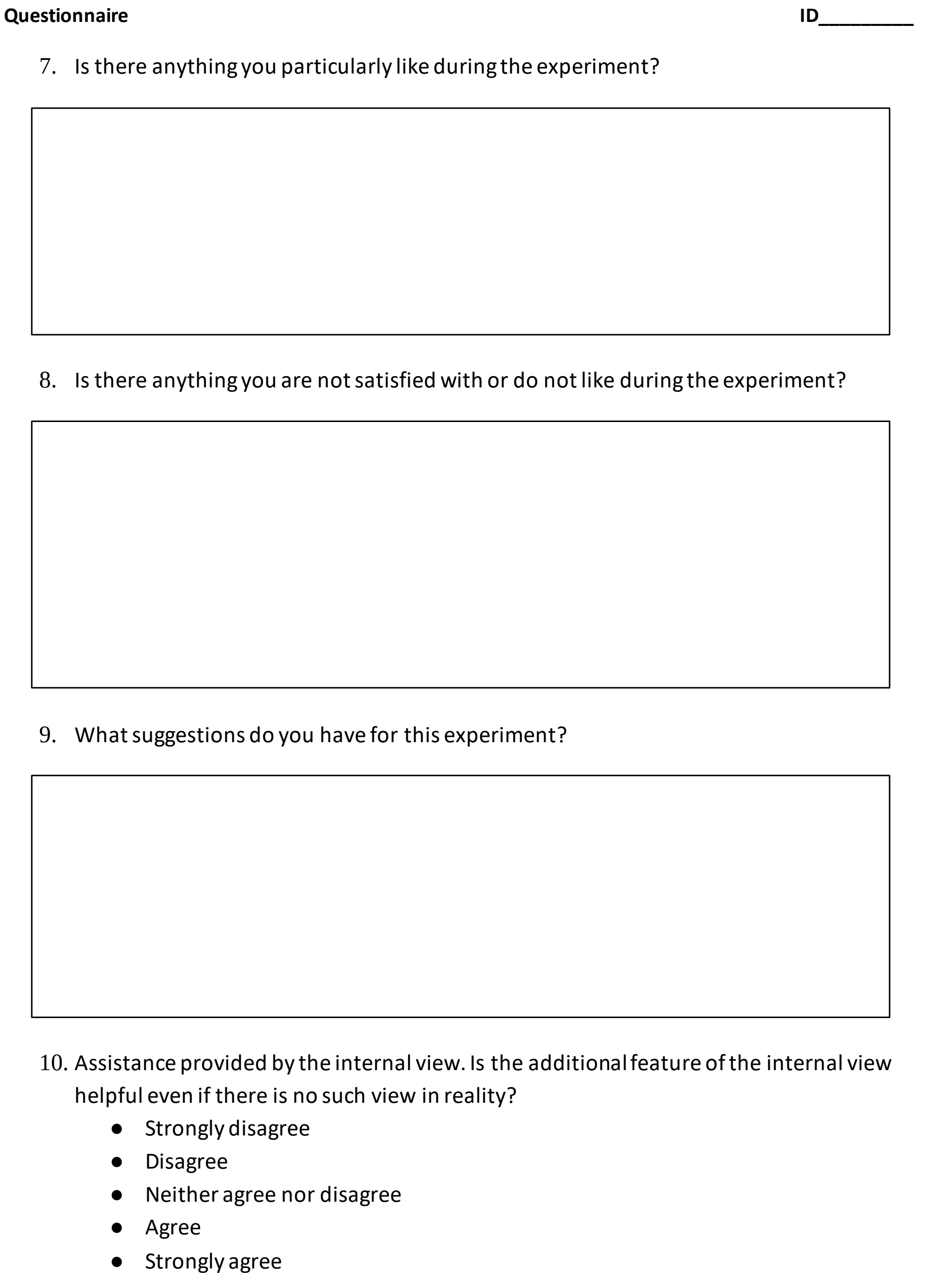
Appendix B

Questionnaire for user study

Questionnaire **ID_________**

Questionnaire on new technologies in advanced vessel navigation (2 pages)

1. Wearing glasses:
   - ☐ yes
   - ☐ no
2. Handedness:
   - ☐ right hand
   - ☐ left hand
3. Experience in Hololens:
   - ☐ novice
   - ☐ intermediate
   - ☐ proficient
4. Experience in gaming:
   - ☐ novice
   - ☐ intermediate
   - ☐ proficient
5. Experience in steerable catheters:
   - ☐ novice
   - ☐ intermediate
   - ☐ proficient
6. Please sort the three modes from highest to lowest according to your preference. Why?



Questionnaire **ID**_________

7. Is there anything you particularly like during the experiment?

8. Is there anything you are not satisfied with or do not like during the experiment?

9. What suggestions do you have for this experiment?

10. Assistance provided by the internal view. Is the additional feature of the internal view helpful even if there is no such view in reality?
    - Strongly disagree
    - Disagree
    - Neither agree nor disagree
    - Agree
    - Strongly agree

Figure B.1: Custom-designed questionnaire employed for the user study presented in Chapter 3.

Bibliography

[1] Chad J Davis. A history of endoscopic surgery, 1992.

[2] R Vecchio, BV MacFayden, and F Palazzo. History of laparoscopic surgery. *Panminerva medica*, 42(1):87–90, 2000.

[3] Nathaniel J Soper, Lee L Swanström, and Steve Eubanks. *Mastery of endoscopic and laparoscopic surgery.* Lippincott Williams & Wilkins, 2008.

[4] Joris EN Jaspers, Paul Breedveld, Just L Herder, and Cornelis A Grimbergen. Camera and instrument holders and their clinical value in minimally invasive surgery. *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, 14(3): 145–152, 2004.

[5] Cinzia Freschi, Vincenzo Ferrari, Franca Melfi, Mauro Ferrari, Franco Mosca, and Alfred Cuschieri. Technical review of the da vinci surgical telemanipulator. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 9(4):396–406, 2013.

[6] Diana M Greenfield and John A Snowden. Cardiovascular diseases and metabolic syndrome. *The EBMT Handbook: Hematopoietic Stem Cell Transplantation and Cellular Therapies*, pages 415–420, 2019.

[7] Sally Robinson. Cardiovascular disease. In *Priorities for Health Promotion and Public Health*, pages 355–393. Routledge, 2021.

[8] Shanthi Mendis, Pekka Puska, B editors Norrving, World Health Organization, et al. *Global atlas on cardiovascular disease prevention and control.* World Health Organization, 2011.

[9] II Abubakar, Taavi Tillmann, and Amitava Banerjee. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the global burden of disease study 2013. *Lancet*, 385(9963):117–171, 2015.

[10] Sudheeran Kannoth, Rajesh Iyer, Sanjeev V Thomas, Sunil V Furtado, BJ Rajesh, C Kesavadas, VV Radhakrishnan, and PS Sarma. Intracranial

infectious aneurysm: presentation, management and outcome. *Journal of the neurological sciences*, 256(1-2):3–9, 2007.

[11] Rulon L Hardman, Omid Jazaeri, J Yi, M Smith, and Rajan Gupta. Overview of classification systems in peripheral artery disease. In *Seminars in interventional radiology*, volume 31, pages 378–388. Thieme Medical Publishers, 2014.

[12] Kayla B Stefanidis, Christopher D Askew, Kim Greaves, and Mathew J Summers. The effect of non-stroke cardiovascular disease states on risk for cognitive decline and dementia: a systematic and meta-analytic review. *Neuropsychology review*, 28:1–15, 2018.

[13] Renata Komalasari, Maria M Yoche, et al. Quality of life of people with cardiovascular disease: a descriptive study. *Asian/Pacific Island Nursing Journal*, 4(2):92, 2019.

[14] Tamás Deme. European research project tests cardiac ct as an alternative to cardiac catheterization. `https://semmelweis.hu/english/2022/03/european-research-project-tests-cardiac-ct-as-an-alternative-to-cardiac-catheterization/`, 2022. Accessed: 2023-03-30.

[15] Steven L Goldberg, Richard Renslo, Robert Sinow, and William J French. Learning curve in the use of the radial artery as vascular access in the performance of percutaneous transluminal coronary angioplasty. *Catheterization and cardiovascular diagnosis*, 44(2):147–152, 1998.

[16] AdventHealth Gordon Cath Lab. AdventHealth Gordon Makes Significant Advances in Cardiology Care with 24/7 PCI and STEMI Services. `https://www.adventhealth.com/hospital/adventhealth-gordon/news/adventhealth-gordon-makes-significant-advances-cardiology-care-24-7-pci-and-stemi-services`, 2022.

[17] Hisato Takagi, Tomo Ando, Takuya Umemoto, ALICE (All-Literature Investigation of Cardiovascular Evidence) Group, et al. A review of comparative studies of mitraclip versus surgical repair for mitral regurgitation. *International journal of cardiology*, 228:289–294, 2017.

[18] Nicholas A Oh, Polydoros N Kampaktsis, Michele Gallo, Alvise Guariento, Viktoria Weixler, Steven J Staffa, Dimitrios V Avgerinos, Andrea Colli, and Ilias P Doulamis. An updated meta-analysis of mitraclip versus surgery for mitral regurgitation. *Annals of Cardiothoracic Surgery*, 10(1):1, 2021.

[19] Kalpa De Silva, Aung Myat, Julian Strange, and Giora Weisz. Iterative improvement and marginal gains in coronary revascularisation: is robot-assisted percutaneous coronary intervention the new hope? *Interventional Cardiology Review*, 15, 2020.

[20] Yu Song, Shuxiang Guo, Xuanchun Yin, Linshuai Zhang, Hideyuki Hirata, Hidenori Ishihara, and Takashi Tamiya. Performance evaluation of a robot-assisted catheter operating system with haptic feedback. *Biomedical microdevices*, 20:1–16, 2018.

[21] Yu Song, Shuxiang Guo, Linshuai Zhang, and Miao Yu. Haptic feedback in robot-assisted endovascular catheterization. In *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 404–409. IEEE, 2017.

[22] Jaehong Woo, Hwa-Seob Song, Hyo-Jeong Cha, and Byung-Ju Yi. Advantage of steerable catheter and haptic feedback for a 5-dof vascular intervention robot system. *Applied Sciences*, 9(20):4305, 2019.

[23] Leonard Berlin. Radiation-induced skin injuries and fluoroscopy. *American Journal of Roentgenology*, 177(1):21–25, 2001.

[24] Marian Limbacher, Pamela S Douglas, Guido Germano, et al. Radiation safety in the practice of cardiology. *Journal of the American College of Cardiology*, 31(4):892–915, 1998.

[25] Guang-Zhong Yang, James Cambias, Kevin Cleary, Eric Daimler, James Drake, Pierre E Dupont, Nobuhiko Hata, Peter Kazanzides, Sylvain Martel, Rajni V Patel, et al. Medical robotics—regulatory, ethical, and legal considerations for increasing levels of autonomy, 2017.

[26] CV Riga, CD Bicknell, R Sidhu, F Cochennec, P Normahani, P Chadha, E Kashef, M Hamady, and NJW Cheshire. Advanced catheter technology: is this the answer to overcoming the long learning curve in complex endovascular procedures. *European Journal of Vascular and Endovascular Surgery*, 42(4): 531–538, 2011.

[27] Arthur Fourcade and Roman Hossein Khonsari. Deep learning in medical image analysis: A third eye for doctors. *Journal of stomatology, oral and maxillofacial surgery*, 120(4):279–288, 2019.

[28] Saima Safdar, Saad Zafar, Nadeem Zafar, and Naurin Farooq Khan. Machine learning based decision support systems (dss) for heart disease diagnosis: a review. *Artificial Intelligence Review*, 50:597–623, 2018.

[29] Alexandru Nicolae, Mark Semple, Lin Lu, Mackenzie Smith, Hans Chung, Andrew Loblaw, Gerard Morton, Lucas Castro Mendez, Chia-Lin Tseng, Melanie Davidson, et al. Conventional vs machine learning–based treatment planning in prostate brachytherapy: results of a phase i randomized controlled trial. *Brachytherapy*, 19(4):470–476, 2020.

[30] Benoit Rosa, Alain Devreker, Herbert De Praetere, Caspar Gruijthuijsen, Sergio Portoles-Diez, Andy Gijbels, Dominiek Reynaerts, Paul Herijgers, Jos

Vander Sloten, and Emmanuel Vander Poorten. Intuitive teleoperation of active catheters for endovascular surgery. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2617–2624. IEEE, 2015.

[31] Di Wu, Mouloud Ourak, Mirza Awais Ahmad, Gianni Borghesan, Jenny Dankelman, and Emmanuel Vander Poorten. Feasibility of using a long short-term memory network for robotic catheter control. In *10 th Conference on New Technologies for Computer and Robot Assisted Surgery, Date: 2020/09/28-2020/09/30, Location: Barcelona, Spain*, pages 68–69, 2020.

[32] Sidrah Khan and Allan Tsung. Aso author reflections: The evolution of minimally invasive liver surgery and the future with robotics. *Annals of surgical oncology*, 25(3):786–787, 2018.

[33] B Jaffray. Minimally invasive surgery. *Archives of disease in childhood*, 90 (5):537–542, 2005.

[34] Xiaohua Hu, Ang Chen, Yigang Luo, Chris Zhang, and Edwin Zhang. Steerable catheters for minimally invasive surgery: a review and future directions. *Computer Assisted Surgery*, 23(1):21–41, 2018.

[35] David B Camarillo, Christopher F Milne, Christopher R Carlson, Michael R Zinn, and J Kenneth Salisbury. Mechanics modeling of tendon-driven continuum manipulators. *IEEE Transactions on Robotics*, 24(6):1262–1273, 2008.

[36] Pierre E Dupont, Jesse Lock, Brandon Itkowitz, and Evan Butler. Design and control of concentric-tube robots. *IEEE Transactions on Robotics*, 26(2): 209–225, 2009.

[37] Pierre Berthet-Rayne, Gauthier Gras, Konrad Leibrandt, Piyamate Wisanuvej, Andreas Schmitz, Carlo A Seneci, and Guang-Zhong Yang. The i 2 snake robotic platform for endoscopic surgery. *Annals of Biomedical Engineering*, 46(10):1663–1675, 2018.

[38] Thomas Watts, Riccardo Secoli, and Ferdinando Rodriguez y Baena. A mechanics-based model for 3-d steering of programmable bevel-tip needles. *IEEE Transactions on Robotics*, 35(2):371–386, 2018.

[39] Jessica Burgner-Kahrs, D Caleb Rucker, and Howie Choset. Continuum robots for medical applications: A survey. *IEEE transactions on Robotics*, 31 (6):1261–1280, 2015.

[40] Valentina Vitiello, Su-Lin Lee, Thomas P Cundy, and Guang-Zhong Yang. Emerging robotic platforms for minimally invasive surgery. *IEEE Reviews in Biomedical Engineering*, 6:111–126, 2012.

[41] Olatunji Mumini Omisore, Shipeng Han, Jing Xiong, Hui Li, Zheng Li, and Lei Wang. A review on flexible robotic systems for minimally invasive surgery. *IEEE transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[42] Tomas da Veiga, James H Chandler, Peter Lloyd, Giovanni Pittiglio, Nathan J Wilkinson, Ali K Hoshiar, Russell A Harris, and Pietro Valdastri. Challenges of continuum robots in clinical context: A review. *Progress in Biomedical Engineering*, 2(3):032003, 2020.

[43] George P Moustris, Savvas C Hiridis, Kyriakos M Deliparaschos, and Konstantinos M Konstantinidis. Evolution of autonomous and semi-autonomous robotic surgical systems: a review of the literature. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7 (4):375–392, 2011.

[44] Shahriar Sefati, Cong Gao, Iulian Iordachita, Russell H Taylor, and Mehran Armand. Data-driven shape sensing of a surgical continuum manipulator using an uncalibrated fiber bragg grating sensor. *IEEE Sensors Journal*, 21 (3):3066–3076, 2020.

[45] Jie Chen and Henry YK Lau. Transferring autonomous reaching and targeting behaviors for cable-driven robots in minimally invasive surgery. In *2016 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, pages 79–84. IEEE, 2016.

[46] Sylvain Calinon, Danilo Bruno, Milad S Malekzadeh, Thrishantha Nanayakkara, and Darwin G Caldwell. Human–robot skills transfer interfaces for a flexible surgical robot. *Computer Methods and Programs in Biomedicine*, 116(2):81–96, 2014.

[47] Wenqiang Chi, Jindong Liu, Hedyeh Rafii-Tari, Celia Riga, Colin Bicknell, and Guang-Zhong Yang. Learning-based endovascular navigation through the use of non-rigid registration for collaborative robotic catheterization. *International Journal of Computer Assisted Radiology and Surgery*, 13(6): 855–864, 2018.

[48] Yohannes Kassahun, Bingbin Yu, Abraham Temesgen Tibebu, Danail Stoyanov, Stamatia Giannarou, Jan Hendrik Metzen, and Emmanuel Vander Poorten. Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions. *The International Journal for Computer Assisted Radiology and Surgery*, 11(4):553–568, 2016.

[49] Trishan Panch, Peter Szolovits, and Rifat Atun. Artificial intelligence, machine learning and health systems. *Journal of Global Health*, 8(2), 2018.

[50] Andrew L Beam and Isaac S Kohane. Big data and machine learning in health care. *Jama*, 319(13):1317–1318, 2018.

[51] Danilo Bruno, Sylvain Calinon, and Darwin G Caldwell. Null space redundancy learning for a flexible surgical robot. In *2014 International Conference on Robotics and Automation (ICRA)*, pages 2443–2448. IEEE, 2014.

[52] Wenqiang Chi, Giulio Dagnino, Trevor MY Kwok, Anh Nguyen, Dennis Kundrat, Mohamed EMK Abdelaziz, Celia Riga, Colin Bicknell, and Guang-Zhong Yang. Collaborative robot-assisted endovascular catheterization with generative adversarial imitation learning. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2414–2420. IEEE, 2020.

[53] Hedyeh Rafii-Tari, Jindong Liu, Christopher J Payne, Colin Bicknell, and Guang-Zhong Yang. Hierarchical hmm based learning of navigation primitives for cooperative robotic endovascular catheterization. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part I 17*, pages 496–503. Springer, 2014.

[54] Xiaoyu Tan, Yonggu Lee, Chin-Boon Chng, Kah-Bin Lim, and Chee-Kong Chui. Robot-assisted flexible needle insertion using universal distributional deep reinforcement learning. *The International Journal for Computer Assisted Radiology and Surgery*, 15(2):341–349, 2020.

[55] Hedyeh Rafii-Tari, Jindong Liu, Su-Lin Lee, Colin Bicknell, and Guang-Zhong Yang. Learning-based modeling of endovascular navigation for collaborative robotic catheterization. In *International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 369–377. Springer, 2013.

[56] Emmanuel Vander Poorten, P Tran, Alain Devreker, Caspar Gruijthuijsen, Sergio Portoles-Diez, Gabrijel Smoljkic, Vule Strbac, Nele Famaey, Dominiek Reynaerts, Jos Vander Sloten, et al. Cognitive autonomous catheters operating in dynamic environments. *Journal of Medical Robotics Research*, 1 (03):1640011, 2016.

[57] Hedyeh Rafii-Tari, Christopher J Payne, and Guang-Zhong Yang. Current and emerging robot-assisted endovascular catheterization technologies: a review. *Annals of Biomedical Engineering*, 42(4):697–715, 2014.

[58] Daekyum Kim, Sang-Hun Kim, Taekyoung Kim, Brian Byunghyun Kang, Minhyuk Lee, Wookeun Park, Subyeong Ku, DongWook Kim, Junghan Kwon, Hochang Lee, et al. Review of machine learning methods in soft robotics. *Plos one*, 16(2):e0246102, 2021.

[59] Ethem Alpaydin. *Introduction to machine learning.* MIT press, 2020.

[60] Kevin P Murphy. *Machine learning: a probabilistic perspective.* MIT press, 2012.

[61] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[62] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[63] Reinhard Grassmann, Vincent Modes, and Jessica Burgner-Kahrs. Learning the forward and inverse kinematics of a 6-dof concentric tube continuum robot in se (3). In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5125–5132. IEEE, 2018.

[64] Wenjun Xu, Jie Chen, Henry YK Lau, and Hongliang Ren. Data-driven methods towards learning the highly nonlinear inverse kinematics of tendon-driven surgical manipulators. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(3):e1774, 2017.

[65] Chaoyang Shi, Xiongbiao Luo, Peng Qi, Tianliang Li, Shuang Song, Zoran Najdovski, Toshio Fukuda, and Hongliang Ren. Shape sensing techniques for continuum robots in minimally invasive surgery: A survey. *IEEE Transactions on Biomedical Engineering*, 64(8):1665–1678, 2016.

[66] Jie Chen, Henry YK Lau, Wenjun Xu, and Hongliang Ren. Towards transferring skills to flexible surgical robots with programming by demonstration and reinforcement learning. In *2016 Eighth International Conference on Advanced Computing Intelligence (ICACI)*, pages 378–384. IEEE, 2016.

[67] Ling Li, Xiaojian Li, Shanlin Yang, Shuai Ding, Alireza Jolfaei, and Xi Zheng. Unsupervised-learning-based continuous depth and motion estimation with monocular endoscopy for virtual reality minimally invasive surgery. *IEEE Transactions on Industrial Informatics*, 17(6):3920–3928, 2020.

[68] Alice Segato, Luca Sestini, Antonella Castellano, and Elena De Momi. Ga3c reinforcement learning for surgical steerable catheter path planning. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2429–2435. IEEE, 2020.

[69] Hyeonseok You, EunKyung Bae, Youngjin Moon, Jihoon Kweon, and Jaesoon Choi. Automatic control of cardiac ablation catheter with deep reinforcement learning method. *Journal of Mechanical Science and Technology*, 33(11): 5415–5423, 2019.

[70] Niall O'Mahony, Sean Campbell, Anderson Carvalho, Suman Harapanahalli, Gustavo Velasco Hernandez, Lenka Krpalkova, Daniel Riordan, and Joseph Walsh. Deep learning vs. traditional computer vision. In *Science and information conference*, pages 128–144. Springer, 2019.

[71] Dr Jennifer S Raj and J Vijitha Ananthi. Recurrent neural networks and nonlinear prediction in support vector machines. *Journal of Soft Computing Paradigm (JSCP)*, 1(01):33–40, 2019.

[72] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. A survey of methods for explaining black box models. *ACM comput. surv. (CSUR)*, 51(5):1–42, 2018.

[73] Hsi Guang Sung. *Gaussian mixture regression and classification.* Rice University, 2004.

[74] Hongqiang Wang, Jie Chen, Henry YK Lau, and Hongliang Ren. Motion planning based on learning from demonstration for multiple-segment flexible soft robots actuated by electroactive polymers. *IEEE Robotics and Automation Letters*, 1(1):391–398, 2016.

[75] Bingbin Yu, José de Gea Fernández, and Tao Tan. Probabilistic kinematic model of a robotic catheter for 3D position control. *Soft Robotics*, 6(2): 184–194, 2019.

[76] Jie Chen and Henry YK Lau. Learning the inverse kinematics of tendon-driven soft manipulators with k-nearest neighbors regression and gaussian mixture regression. In *2016 2nd International Conference on Control, Automation and Robotics (ICCAR)*, pages 103–107. IEEE, 2016.

[77] Mohammad Jolaei, Amir Hooshiar, Javad Dargahi, and Muthukumaran Packirisamy. Toward task autonomy in robotic cardiac ablation: Learning-based kinematic control of soft tendon-driven catheters. *Soft Robotics*, 8(3): 340–351, 2021.

[78] Georgios Fagogenis, Margherita Mencattelli, Zurab Machaidze, Benoit Rosa, Karl Price, F Wu, V Weixler, Mossab Saeed, John E Mayer, and Pierre E Dupont. Autonomous robotic intracardiac catheter navigation using haptic vision. *Science Robotics*, 4(29):eaaw1977, 2019.

[79] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[80] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The Handbook of Brain Theory and Neural Networks*, 3361(10):1995, 1995.

[81] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

[82] Yu Emma Wang, Gu-Yeon Wei, and David Brooks. Benchmarking TPU, GPU, and CPU platforms for deep learning. *arXiv preprint arXiv:1907.10701*, 2019.

[83] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32:8026–8037, 2019.

[84] Sara Gerke, Timo Minssen, and Glenn Cohen. Ethical and legal challenges of artificial intelligence-driven healthcare. In *Artificial intelligence in healthcare*, pages 295–336. Elsevier, 2020.

[85] Olatunji Mumini Omisore, Shipeng Han, Tao Zhou, Yousef Al-Handarish, Wenjing Du, Kamen Ivanov, and Lei Wang. Learning-based parameter estimation for hysteresis modeling in robotic catheterization. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5399–5402. IEEE, 2019.

[86] Thomas George Thuruthel, Egidio Falotico, Matteo Cianchetti, and Cecilia Laschi. Learning global inverse kinematics solutions for a continuum robot. In *Symposium on Robot Design, Dynamics and Control*, pages 47–54. Springer, 2016.

[87] Jake Sganga, David Eng, Chauncey Graetzel, and David B Camarillo. Autonomous driving in the lung using deep learning for localization. *arXiv preprint arXiv:1907.08136*, 2019.

[88] Mirza Awais Ahmad, Mouloud Ourak, Caspar Gruijthuijsen, Jan Deprest, Tom Vercauteren, and Emmanuel Vander Poorten. Deep learning-based monocular placental pose estimation: towards collaborative robotics in fetoscopy. *The International Journal for Computer Assisted Radiology and Surgery*, 15(9):1561–1571, 2020.

[89] Xiaoguo Li, Anthony Meng Huat Tiong, Lin Cao, Wenjie Lai, Phuoc Thien Phan, and Soo Jay Phee. Deep learning for haptic feedback of flexible endoscopic robot without prior knowledge on sheath configuration. *International Journal of Mechanical Sciences*, 163:105129, 2019.

[90] Md Zahangir Alom, Tarek M Taha, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C Van Esesn, Abdul A S Awwal, and Vijayan K Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*, 2018.

[91] Di Wu, Yao Zhang, Mouloud Ourak, Kenan Niu, Jenny Dankelman, and Emmanuel Vander Poorten. Hysteresis modeling of robotic catheters based on long short-term memory network for improved environment reconstruction. *IEEE Robotics and Automation Letters*, 6(2):2106–2113, 2021.

[92] Xiaoguo Li, Lin Cao, Anthony Meng Huat Tiong, Phuoc Thien Phan, and Soo Jay Phee. Distal-end force prediction of tendon-sheath mechanisms

for flexible endoscopic surgical robots using deep learning. *Mechanism and Machine Theory*, 134:323–337, 2019.

[93] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[94] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, 2018.

[95] Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, et al. Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv:1912.06680*, 2019.

[96] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[97] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. Deterministic policy gradient algorithms. In *International Conference on Machine Learning*, pages 387–395. PMLR, 2014.

[98] Keshav Iyengar, George Dwyer, and Danail Stoyanov. Investigating exploration for deep reinforcement learning of concentric tube robot control. *The International Journal for Computer Assisted Radiology and Surgery*, 15 (7):1157–1165, 2020.

[99] Tobias Behr, Tim Philipp Pusch, Marius Siegfarth, Dominik Hüsener, Tobias Mörschel, and Lennart Karstensen. Deep reinforcement learning for the navigation of neurovascular catheters. *Current Directions in Biomedical Engineering*, 5(1):5–8, 2019.

[100] Sreeshankar Satheeshbabu, Naveen Kumar Uppalapati, Girish Chowdhary, and Girish Krishnan. Open loop position control of soft continuum arm using deep reinforcement learning. In *2019 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5133–5139. IEEE, 2019.

[101] Michael E Rose and John R Kitchin. pybliometrics: Scriptable bibliometrics using a python interface to scopus. *SoftwareX*, 10:100263, 2019.

[102] Nao Ito, Hiroshi Kawahira, Hirotaka Nakashima, Masaya Uesato, Hideaki Miyauchi, and Hisahiro Matsubara. Endoscopic diagnostic support system for ct1b colorectal cancer using deep learning. *Oncology*, 96(1):44–50, 2019.

[103] Patrick Brandao, Odysseas Zisimopoulos, Evangelos Mazomenos, Gastone Ciuti, Jorge Bernal, Marco Visentini-Scarzanella, Arianna Menciassi, Paolo Dario, Anastasios Koulaouzidis, Alberto Arezzo, et al. Towards a computed-aided diagnosis system in colonoscopy: automatic polyp segmentation using convolution neural networks. *Journal of Medical Robotics Research*, 3(02): 1840002, 2018.

[104] Ling Li, Xiaojian Li, Shanlin Yang, Shuai Ding, Alireza Jolfaei, and Xi Zheng. Unsupervised-learning-based continuous depth and motion estimation with monocular endoscopy for virtual reality minimally invasive surgery. *IEEE Transactions on Industrial Informatics*, 17(6):3920–3928, 2020.

[105] Weibing Li, Philip Wai Yan Chiu, and Zheng Li. An accelerated finite-time convergent neural network for visual servoing of a flexible surgical endoscope with physical and rcm constraints. *IEEE transactions on Neural Networks and Learning Systems*, 31(12):5272–5284, 2020.

[106] Weibing Li, Chengzhi Song, and Zheng Li. An accelerated recurrent neural network for visual servo control of a robotic flexible endoscope with joint limit constraint. *IEEE Transactions on Industrial Electronics*, 67(12):10787–10797, 2019.

[107] Shahriar Sefati, Shahin Sefati, Iulian Iordachita, Russell H Taylor, and Mehran Armand. Learning to detect collisions for continuum manipulators without a prior model. In *International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 182–190. Springer, 2019.

[108] Heiko Donat, Sven Lilge, Jessica Burgner-Kahrs, and Jochen J Steil. Estimating tip contact forces for concentric tube continuum robots based on backbone deflection. *IEEE Transactions on Medical Robotics and Bionics*, 2 (4):619–630, 2020.

[109] Seunghyuk Cho, Minsoo Koo, Dongwoo Kim, Juyong Lee, Yeonwoo Jung, Kibyung Nam, and Changmo Hwang. Roughly collected dataset for contact force sensing catheter. *arXiv preprint arXiv:2102.01932*, 2021.

[110] Alan Kuntz, Armaan Sethi, Robert J Webster, and Ron Alterovitz. Learning the complete shape of concentric tube robots. *IEEE Transactions on Medical Robotics and Bionics*, 2(2):140–147, 2020.

[111] Connor Watson and Tania K Morimoto. Permanent magnet-based localization for growing robots in medical applications. *IEEE Robotics and Automation Letters*, 5(2):2666–2673, 2020.

[112] Wenqiang Chi, Jindong Liu, Mohamed Emk Abdelaziz, Giulio Dagnino, Celia Riga, Colin Bicknell, and Guang-Zhong Yang. Trajectory optimization of robot-assisted endovascular catheterization with reinforcement learning. In

2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 3875–3881. IEEE, 2018.

[113] Kit-Hang Lee, Denny KC Fu, Martin CW Leong, Marco Chow, Hing-Choi Fu, Kaspar Althoefer, Kam Yim Sze, Chung-Kwong Yeung, and Ka-Wai Kwok. Nonparametric online learning control for soft continuum robot: An enabling technique for effective endoscopic navigation. *Soft robotics*, 4(4): 324–337, 2017.

[114] Olatunji Mumini Omisore, Toluwanimi Akinyemi, Wenke Duan, Wenjing Du, and Lei Wang. A novel sample-efficient deep reinforcement learning with episodic policy transfer for pid-based control in cardiac catheterization robots. *arXiv e-prints*, pages arXiv–2110, 2021.

[115] M Goharimanesh, Ali Mehrkish, and Farrokh Janabi-Sharifi. A fuzzy reinforcement learning approach for continuum robot control. *Journal of Intelligent & Robotic Systems*, 100(3):809–826, 2020.

[116] Rafael Aleluia Porto, Florent Nageotte, Philippe Zanne, and Michel de Mathelin. Position control of medical cable-driven flexible instruments by combining machine learning and kinematic analysis. In *2019 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7913–7919. IEEE, 2019.

[117] Wei Jiang, Tao Yu, Xiao He, Yongming Yang, Zhidong Wang, and Hao Liu. Data-driven modeling the nonlinear backlash of steerable endoscope under a large deflection cannulation in ercp surgery. In *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pages 39–44. IEEE, 2021.

[118] Francesco Cursi, Valerio Modugno, Leonardo Lanari, Giuseppe Oriolo, and Petar Kormushev. Bayesian neural network modeling and hierarchical mpc for a tendon-driven surgical robot with uncertainty minimization. *IEEE Robotics and Automation Letters*, 6(2):2642–2649, 2021.

[119] Weibang Bai, Francesco Cursi, Xiaotong Guo, Baoru Huang, Benny Lo, Guang-Zhong Yang, and Eric M Yeatman. Task-based lstm kinematic modeling for a tendon-driven flexible surgical robot. *IEEE Transactions on Medical Robotics and Bionics*, 4(2):339–342, 2021.

[120] Yan Zhao, Shuxiang Guo, Yuxin Wang, Jinxin Cui, Youchun Ma, Yuwen Zeng, Xinke Liu, Yuhua Jiang, Youxinag Li, Liwei Shi, et al. A CNN-based prototype method of unstructured surgical state perception and navigation for an endovascular surgery robot. *Medical & Biological Engineering & Computing*, 57(9):1875–1887, 2019.

[121] Jake Sganga. *Autonomous Navigation of a Flexible Surgical Robot in the Lungs*. Stanford University, 2019.

[122] Milad S Malekzadeh, Sylvain Calinon, Danilo Bruno, and Darwin G Caldwell. A skill transfer approach for continuum robots—imitation of octopus reaching motion with the stiff-flop robot. In *2014 AAAI Fall Symposium Series*, 2014.

[123] Milad S Malekzadeh, Danilo Bruno, Sylvain Calinon, Thrishantha Nanayakkara, and Darwin G Caldwell. Skills transfer across dissimilar robots by learning context-dependent rewards. In *2013 IEEE/RSJ International Conference on Intell. Robots and Syst.*, pages 1746–1751. IEEE, 2013.

[124] Ryan L Truby, Cosimo Della Santina, and Daniela Rus. Distributed proprioception of 3D configuration in soft, sensorized robots via deep learning. *IEEE Robotics and Automation Letters*, 5(2):3299–3306, 2020.

[125] Samaneh Manavi, Tatiana Renna, Antal Horvath, Sara Freund, Azhar Zam, Georg Rauter, Wolfgang Schade, and Philippe C Cattin. Using supervised deep-learning to model edge-fbg shape sensors: a feasibility study. In *Optical Sensors 2021*, volume 11772, pages 79–88. SPIE, 2021.

[126] Xuan Thao Ha, Di Wu, Mouloud Ourak, Gianni Borghesan, Jenny Dankelman, Arianna Menciassi, and Emmanuel Vander Poorten. Shape sensing of flexible robots based on deep learning. *IEEE Transactions on Robotics*, 2022.

[127] Cosimo Della Santina, Ryan Landon Truby, and Daniela Rus. Data–driven disturbance observers for estimating external forces on soft robots. *IEEE Robotics and Automation Letters*, 5(4):5717–5724, 2020.

[128] Xuan Thao Ha, Di Wu, Chun-Feng Lai, Mouloud Ourak, Gianni Borghesan, Arianna Menciassi, and Emmanuel Vander Poorten. Contact localization of continuum and flexible robot using data-driven approach. *IEEE Robotics and Automation Letters*, 7(3):6910–6917, 2022. doi: 10.1109/LRA.2022.3176723.

[129] Xin Yu, Rajinder P Singh-Moon, and Christine P Hendon. Real-time assessment of catheter contact and orientation using an integrated optical coherence tomography cardiac ablation catheter. *Applied optics*, 58(14): 3823–3829, 2019.

[130] C Bergeles, FY Lin, and GZ Yang. Concentric tube robot kinematics using neural networks. In *Hamlyn Symposium on Medical Robotics*, pages 13–14, 2015.

[131] W Li, F Cursi, W Bai, and G Mylonas. Poster: Learning-based modeling and control of micro-iges surgical robot. In *Hamlyn Symposium on Medical Robotics*, June 2021.

[132] Donghoon Baek, Ju-Hwan Seo, Joonhwan Kim, and Dong-Soo Kwon. Hysteresis compensator with learning-based pose estimation for a flexible endoscopic surgery robot. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2983–2989. IEEE, 2019.

[133] Baek, Donghoon and Seo, Ju-Hwan and Kim, Joonhwan and Kwon, Dong-Soo. Hysteresis compensator with learning-based hybrid joint angle estimation for flexible surgery robots. *IEEE Robotics and Automation Letters*, 5(4): 6837–6844, 2020.

[134] Olatunji Mumini Omisore, Shi Peng Han, Ling Xue Ren, Guo Shuai Wang, Fang Liang Ou, Hui Li, and Lei Wang. Towards characterization and adaptive compensation of backlash in a novel robotic catheter system for cardiovascular interventions. *IEEE Transactions on Biomedical Circuits and Systems*, 12 (4):824–838, 2018.

[135] Jie Chen and Henry Lau. Policy gradient-based inverse kinematics refinement for tendon-driven serpentine surgical manipulator. *International Journal of Robotics and Automation*, 34(3), 2019.

[136] Mark Edgington, Yohannes Kassahun, and Frank Kirchner. Using joint probability densities for simultaneous learning of forward and inverse models. In *IEEE IROS International Workshop on Evolutionary and Reinforcement Learning for Autonomous Robot Systems, NT Siebel and J. Pauli, Eds*, volume 10, pages 19–22, 2009.

[137] Michele Giorelli, Federico Renda, Gabriele Ferri, and Cecilia Laschi. A feed-forward neural network learning the inverse kinetics of a soft cable-driven manipulator moving in three-dimensional space. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5033–5039, 2013. doi: 10.1109/IROS.2013.6697084.

[138] Di Wu, Xuan Thao Ha, Yao Zhang, Mouloud Ourak, Gianni Borghesan, Kenan Niu, Fabian Trauzettel, Jenny Dankelman, Arianna Menciassi, and Emmanuel Vander Poorten. Deep-learning-based compliant motion control of a pneumatically-driven robotic catheter. *IEEE Robotics and Automation Letters*, 7(4):8853–8860, 2022.

[139] Saeid Shakiba, Mouloud Ourak, Emmanuel Vander Poorten, Moosa Ayati, and Aghil Yousefi-Koma. Modeling and compensation of asymmetric rate-dependent hysteresis of a miniature pneumatic artificial muscle-based catheter. *Mechanical System and Signal Processing*, 154:107532, 2021. ISSN 10961216. doi: 10.1016/j.ymssp.2020.107532.

[140] J-SR Jang. Anfis: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665–685, 1993.

[141] Yuxin Wang, Shuxiang Guo, Yan Zhao, Jinxin Cui, and Youchun Ma. A CNNs-based of force and torque identification model for vascular interventional surgery robot. In *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 2291–2296. IEEE, 2019.

[142] Gabriele Trovato, M Shikanai, Genya Ukawa, J Kinoshita, N Murai, JW Lee, Hiroyuki Ishii, Atsuo Takanishi, Kazuo Tanoue, Satoshi Ieiri, et al. Development of a colon endoscope robot that adjusts its locomotion through the use of reinforcement learning. *International Journal of Computer Assisted Radiology and Surgery*, 5(4):317–325, 2010.

[143] Keshav Iyengar and Danail Stoyanov. Deep reinforcement learning for concentric tube robot control with a goal-based curriculum. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1459–1465. IEEE, 2021.

[144] Mohammad Ali Armin, Nick Barnes, Jose Alvarez, Hongdong Li, Florian Grimpen, and Olivier Salvado. Learning camera pose from optical colonoscopy frames through deep convolutional neural network (CNN). In *Computer Assisted and Robotic Endoscopy and Clinical Image-based Procedures*, pages 50–59. Springer, 2017.

[145] Alice Segato, Marco Di Marzo, Sara Zucchelli, Stefano Galvan, Riccardo Secoli, and Elena De Momi. Inverse reinforcement learning intra-operative path planning for steerable needle. *IEEE Transactions on Biomedical Engineering*, 69(6):1995–2005, 2021.

[146] Danilo Bruno, Sylvain Calinon, and G Caldwell, Darwin. Learning autonomous behaviours for the body of a flexible surgical robot. *Autonomous Robots*, 41(2):333–347, 2017.

[147] Wenjun Xu, Jie Chen, Henry YK Lau, and Hongliang Ren. Automate surgical tasks for a flexible serpentine manipulator via learning actuation space trajectory from demonstration. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4406–4413. IEEE, 2016.

[148] Milad S Malekzadeh, Sylvain Calinon, Danilo Bruno, and Darwin G Caldwell. Learning by imitation with the stiff-flop surgical robot: a biomimetic approach inspired by octopus movements. *Robotics and Biomimetics*, 1(1):1–15, 2014.

[149] Abraham Temesgen Tibebu, Bingbin Yu, Yohannes Kassahun, Emmanuel Vander Poorten, and Phuong Toan Tran. Towards autonomous robotic catheter navigation using reinforcement learning. In *4th Joint Workshop on New Technologies for Computer/Robot Assisted Surgery*, pages 163–166, 2014.

[150] Lennart Karstensen, Tobias Behr, Tim Philipp Pusch, Franziska Mathis-Ullrich, and Jan Stallkamp. Autonomous guidewire navigation in a two dimensional vascular phantom. *Current Directions in Biomedical Engineering*, 6(1), 2020.

[151] S Athiniotis, RA Srivatsan, and H Choset. Deep q reinforcement learning for autonomous navigation of surgical snake robot in confined spaces. In *The Hamlyn Symposium on Medical Robotics*, 2019.

[152] Jihoon Kweon, Kyunghwan Kim, Chaehyuk Lee, Hwi Kwon, Jinwoo Park, Kyoseok Song, Young In Kim, Jeeone Park, Inwook Back, Jae-Hyung Roh, et al. Deep reinforcement learning for guidewire navigation in coronary artery phantom. *IEEE Access*, 9:166409–166422, 2021.

[153] Xiaoyu Tan, Pengqian Yu, Kah-Bin Lim, and Chee-Kong Chui. Robust path planning for flexible needle insertion using markov decision processes. *International Journal of Computer Assisted Radiology and Surgery*, 13(9): 1439–1451, 2018.

[154] Yonggu Lee, Xiaoyu Tan, Chin-Boon Chng, and Chee-Kong Chui. Simulation of robot-assisted flexible needle insertion using deep q-network. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 342–346. IEEE, 2019.

[155] Fanxu Meng, Shuxiang Guo, Wei Zhou, and Zhengyang Chen. Evaluation of a reinforcement learning algorithm for vascular intervention surgery. In *2021 IEEE International Conference on Mechatronics and Automation (ICMA)*, pages 1033–1037. IEEE, 2021.

[156] Jean-Claude Latombe. *Robot motion planning.* Springer Science & Business Media, vol. 124, 2012.

[157] Alice Segato, Testa Irene, and Elena DE MOMI. 3D neurosurgical simulator for training robotic steerable catheter agents using generative adversarial imitation learning. In *2020 International Conference on Robotics and Automation.* FRA, 2020.

[158] Sinan Çalışır and Meltem Kurt Pehlivanoğlu. Model-free reinforcement learning algorithms: A survey. In *2019 27th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE, 2019.

[159] Matteo Hessel, Joseph Modayil, Hado Van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, and David Silver. Rainbow: Combining improvements in deep reinforcement learning. In *Thirty-second AAAI Conference on Artificial Intelligence*, 2018.

[160] Jinao Zhang, Yongmin Zhong, and Chengfan Gu. Deformable models for surgical simulation: a survey. *IEEE Reviews in Biomedical Engineering*, 11: 143–164, 2017.

[161] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 737–744. IEEE, 2020.

[162] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural

networks from simulation to the real world. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 23–30. IEEE, 2017.

[163] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3803–3810. IEEE, 2018.

[164] Shan Lin, Fangbo Qin, Yangming Li, Randall A Bly, Kris S Moe, and Blake Hannaford. Lc-gan: Image-to-image translation based on generative adversarial network for endoscopic images. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2914–2920. IEEE, 2020.

[165] Jake Sganga, David Eng, Chauncey Graetzel, and David Camarillo. Offsetnet: Deep learning for localization in the lung using rendered images. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5046–5052. IEEE, 2019.

[166] Tim Miller. Explanation in artificial intelligence: Insights from the social sciences. *Artificial intelligence*, 267:1–38, 2019.

[167] European Parliament and Council of the European Union. Artificial intelligence act: Regulation laying down harmonised rules on artificial intelligence and amending certain union legislative acts. *Proposal for Regulation COM/2021/206 final, Brussels, Belgium*, 2021.

[168] Baowei Fei, Wan Sing Ng, Sunita Chauhan, and Chee Keong Kwoh. The safety issues of medical robotics. *Reliability Engineering & System Safety*, 73 (2):183–192, 2001.

[169] Changliu Liu, Tomer Arnon, Christopher Lazarus, Christopher Strong, Clark Barrett, Mykel J Kochenderfer, et al. Algorithms for verifying deep neural networks. *Found. and Trends® in Optimization*, 4(3-4):244–404, 2021.

[170] Robert H Wortham, Andreas Theodorou, and Joanna J Bryson. Improving robot transparency: Real-time visualisation of robot ai substantially improves understanding in naive observers. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 1424–1431. IEEE, 2017.

[171] Ameya Pore, Davide Corsi, Enrico Marchesini, Diego Dall'Alba, Alicia Casals, Alessandro Farinelli, and Paolo Fiorini. Safe reinforcement learning using formal verification for tissue retraction in autonomous robotic-assisted surgery. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4025–4031. IEEE, 2021.

[172] Javier Garcıa and Fernando Fernández. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research*, 16(1):1437–1480, 2015.

[173] Arjun K Manrai, Birgit H Funke, Heidi L Rehm, Morten S Olesen, Bradley A Maron, Peter Szolovits, David M Margulies, Joseph Loscalzo, and Isaac S Kohane. Genetic misdiagnoses and the potential for health disparities. *New England Journal of Medicine*, 375(7):655–665, 2016.

[174] Anat Reiner Benaim, Ronit Almog, Yuri Gorelik, Irit Hochberg, Laila Nassar, Tanya Mashiach, Mogher Khamaisi, Yael Lurie, Zaher S Azzam, Johad Khoury, et al. Analyzing medical research results based on synthetic data and their relation to real data results: systematic comparison from five observational studies. *JMIR Medical Informatics*, 8(2):e16492, 2020.

[175] Richard J Chen, Ming Y Lu, Tiffany Y Chen, Drew FK Williamson, and Faisal Mahmood. Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, pages 1–5, 2021.

[176] Sana Tonekaboni, Shalmali Joshi, Melissa D McCradden, and Anna Goldenberg. What clinicians want: contextualizing explainable machine learning for clinical end use. In *Machine Learning for Healthcare Conference*, pages 359–380. PMLR, 2019.

[177] Pete S Batra, Matthew W Ryan, Raj Sindwani, and Bradley F Marple. Balloon catheter technology in rhinology: reviewing the evidence. *The Laryngoscope*, 121(1):226–232, 2011.

[178] Bradley C Clark, Kohei Sumihara, Robert McCarter, Charles I Berul, and Jeffrey P Moak. Getting to zero: impact of electroanatomical mapping on fluoroscopy use in pediatric catheter ablation. *Journal of Interventional Cardiac Electrophysiology*, 46:183–189, 2016.

[179] Cosette M Stahl, Quinn C Meisinger, Michael P Andre, Thomas B Kinney, and Isabel G Newton. Radiation risk to the fluoroscopy operator and staff. *American Journal of Roentgenology*, 207(4):737–744, 2016.

[180] Shaodong Li, Zhijiang Du, and Hongjian Yu. A robot-assisted spine surgery system based on intraoperative 2d fluoroscopy navigation. *IEEE Access*, 8: 51786–51802, 2020.

[181] Yanhao Zhang, Liang Zhao, and Shoudong Huang. Aortic 3d deformation reconstruction using 2d x-ray fluoroscopy and 3d pre-operative data for endovascular interventions. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2393–2399. IEEE, 2020.

[182] Pierre Ambrosini, Ihor Smal, Daniel Ruijters, Wiro J Niessen, Adriaan Moelker, and Theo Van Walsum. A hidden markov model for 3d catheter tip

tracking with 2d x-ray catheterization sequence and 3d rotational angiography. *IEEE Transactions on Medical Imaging*, 36(3):757–768, 2016.

[183] Elena K Grant, Joshua P Kanter, Laura J Olivieri, Russell R Cross, Adrienne Campbell-Washburn, Anthony Z Faranesh, Ileen Cronin, Karin S Hamann, Michael L O'Byrne, Michael C Slack, et al. X-ray fused with mri guidance of pre-selected transcatheter congenital heart disease interventions. *Catheterization and Cardiovascular Interventions*, 94(3):399–408, 2019.

[184] Muhammad U Farooq, Atul Khasnis, Arshad Majid, and Mounzer Y Kassab. The role of optical coherence tomography in vascular medicine. *Vascular Medicine*, 14(1):63–71, 2009.

[185] Romaric Loffroy, Nicolas Falvo, Christophe Galland, Léo Fréchier, Frédérik Ledan, Marco Midulla, and Olivier Chevallier. Intravascular ultrasound in the endovascular treatment of patients with peripheral arterial disease: current role and future perspectives. *Frontiers in Cardiovascular Medicine*, 7:551861, 2020.

[186] YingLiang Ma, Rashed Karim, R James Housden, Geert Gijsbers, Roland Bullens, Christopher Aldo Rinaldi, Reza Razavi, Tobias Schaeffter, and Kawal S Rhode. Cardiac unfold: a novel technique for image-guided cardiac catheterization procedures. In *Information Processing in Computer-Assisted Interventions: Third International Conference, IPCAI 2012, Pisa, Italy, June 27, 2012. Proceedings 3*, pages 104–114. Springer, 2012.

[187] Elchanan Bruckheimer, Carmel Rotschild, Tamir Dagan, Gabriel Amir, Aviad Kaufman, Shaul Gelman, and Einat Birk. Computer-generated real-time digital holography: first time use in clinical medical imaging. *European Heart Journal-Cardiovascular Imaging*, 17(8):845–849, 2016.

[188] Catherine Diaz, Michael Walker, Danielle Albers Szafir, and Daniel Szafir. Designing for depth perceptions in augmented reality. In *2017 IEEE international symposium on mixed and augmented reality (ISMAR)*, pages 111–122. IEEE, 2017.

[189] Jun Liu, Subhi J Al'Aref, Gurpreet Singh, Alexandre Caprio, Amir Ali Amiri Moghadam, Sun-Joo Jang, S Chiu Wong, James K Min, Simon Dunham, and Bobak Mosadegh. An augmented reality system for image guidance of transcatheter procedures for structural heart disease. *PloS one*, 14(7): e0219174, 2019.

[190] Maria C Palumbo, Laura Morchi, Valentina Corbetta, Arianna Menciassi, Elena De Momi, Emiliano Votta, and Alberto Redaelli. An easy and user independent augmented reality based navigation system for radiation-free interventional procedure. In *2022 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE, 2022.

[191] Verónica García-Vázquez, Felix Von Haxthausen, Sonja Jäckle, Christian Schumann, Ivo Kuhlemann, Juljan Bouchagiar, Anna-Catharina Höfer, Florian Matysiak, Gereon Hüttmann, Jan Peter Goltz, et al. Navigation and visualisation with hololens in endovascular aortic repair. *Innovative surgical sciences*, 3(3):167–177, 2018.

[192] Cristian A Linte, John Moore, Andrew Wiles, Jennifer Lo, Chris Wedlake, and Terry M Peters. In vitro cardiac catheter navigation via augmented reality surgical guidance. In *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, volume 7261, pages 231–239. SPIE, 2009.

[193] Navid Feizi, Mahdi Tavakoli, Rajni V Patel, and S Farokh Atashzar. Robotics and ai for teleoperation, tele-assessment, and tele-training for surgery in the era of covid-19: existing challenges, and future vision. *Frontiers in Robotics and AI*, 8:610677, 2021.

[194] Kundong Wang, Qingsheng Lu, Bin Chen, Yu Shen, Hongbing Li, Manhua Liu, and Zhuoyan Xu. Endovascular intervention robot with multi-manipulators for surgical procedures: Dexterity, adaptability, and practicability. *Robotics and Computer-Integrated Manufacturing*, 56:75–84, 2019.

[195] David Filgueiras-Rama, Alejandro Estrada, Josh Shachar, Sergio Castrejón, David Doiny, Marta Ortega, Eli Gang, and José L Merino. Remote magnetic navigation for accurate, real-time catheter positioning and ablation in cardiac electrophysiology procedures. *JoVE (Journal of Visualized Experiments)*, (74):e3658, 2013.

[196] Christopher J Payne, Hedyeh Rafii-Tari, and Guang-Zhong Yang. A force feedback system for endovascular catheterisation. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1298–1304. IEEE, 2012.

[197] MA Van Veelen, CJ Snijders, E Van Leeuwen, RHM Goossens, and Geert Kazemier. Improvement of foot pedals used during surgery based on new ergonomic guidelines. *Surgical Endoscopy And Other Interventional Techniques*, 17:1086–1091, 2003.

[198] Pavel Zikmund, Miroslav Macík, and Zdenek Míkovec. Reaction to directional vibrations applied on a joystick. In *New Trends in Civil Aviation*, pages 107–111. CRC Press, 2018.

[199] Daniel Natapov and I Scott MacKenzie. The trackball controller: improving the analog stick. In *Proceedings of the International Academic Conference on the Future of Game Design and Technology*, pages 175–182, 2010.

[200] Alain Devreker, Benoit Rosa, Adrien Desjardins, Erwin J Alles, Luis C Garcia-Peraza, Efthymios Maneas, Danail Stoyanov, Anna L David, Tom Vercauteren, Jan Deprest, et al. Fluidic actuation for intra-operative in situ

imaging. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1415–1421. IEEE, 2015.

[201] Raimund Erbel and Holger Eggebrecht. Aortic dimensions and the risk of dissection. *Heart*, 92(1):137–142, 2006.

[202] Omar Al-Ahmad, Mouloud Ourak, Johan Vlekken, and Emmanuel Vander Poorten. Force control with a novel robotic catheterization system based on braided sleeve grippers. *IEEE Transactions on Medical Robotics and Bionics*, pages 1–1, 2023. doi: 10.1109/TMRB.2023.3291026. DOI: 10.1109/TMRB.2023.3291026.

[203] Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, Andrew Y Ng, et al. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, Japan, 2009.

[204] K Somani Arun, Thomas S Huang, and Steven D Blostein. Least-squares fitting of two 3-d point sets. *IEEE Transactions on pattern analysis and machine intelligence*, (5):698–700, 1987.

[205] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[206] S. S. SHAPIRO and M. B. WILK. An analysis of variance test for normality (complete samples)†. *Biometrika*, 52(3-4):591–611, 12 1965. ISSN 0006-3444. doi: 10.1093/biomet/52.3-4.591.

[207] Patrick E McKight and Julius Najab. Kruskal-wallis test. *The corsini encyclopedia of psychology*, pages 1–1, 2010.

[208] B. L. WELCH. The generalization of 'student's' problem when several different population varlances are involved. *Biometrika*, 34(1-2):28–35, 01 1947. ISSN 0006-3444. doi: 10.1093/biomet/34.1-2.28.

[209] Jerry L Hintze and Ray D Nelson. Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184, 1998.

[210] Stanislao Grazioso, Giuseppe Di Gironimo, and Bruno Siciliano. A geometrically exact model for soft continuum robots: The finite element deformation space formulation. *Soft robotics*, 6(6):790–811, 2019.

[211] Martin Wagner, Sebastian Schafer, Charles Strother, and Charles Mistretta. 4d interventional device reconstruction from biplane fluoroscopy. *Medical physics*, 43(3):1324–1334, 2016.

[212] Shuang Song, Zheng Li, Haoyong Yu, and Hongliang Ren. Electromagnetic positioning for tip tracking and shape sensing of flexible robots. *IEEE Sensors Journal*, 15(8):4565–4575, 2015.

[213] James Avery, Mark Runciman, Ara Darzi, and George P Mylonas. Shape sensing of variable stiffness soft robots using electrical impedance tomography. In *2019 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9066–9072, 2019.

[214] G. Gerboni, A. Diodato, G. Ciuti, M. Cianchetti, and A. Menciassi. Feedback control of soft robot actuators via commercial flex bend sensors. *IEEE/ASME Transactions on Mechatronics*, 22(4):1881–1888, 2017. doi: 10.1109/TMECH.2017.2699677.

[215] Ran Xu, Aaron Yurkewich, and Rajni V Patel. Curvature, torsion, and force sensing in continuum robots using helically wrapped fbg sensors. *IEEE Robotics and Automation Letters*, 1(2):1052–1059, 2016.

[216] Vincent Modes, Tobias Ortmaier, and Jessica Burgner-Kahrs. Shape sensing based on longitudinal strain measurements considering elongation, bending, and twisting. *IEEE Sensors Journal*, 21(5):6712–6723, 2020.

[217] Omar Al-Ahmad, Mouloud Ourak, Jan Van Roosbroeck, Johan Vlekken, and Emmanuel Vander Poorten. Improved fbg-based shape sensing methods for vascular catheterization treatment. *IEEE Robotics and Automation Letters*, 5(3):4687–4694, 2020.

[218] Costanza Culmone, Paul WJ Henselmans, Remi IB van Starkenburg, and Paul Breedveld. Exploring non-assembly 3d printing for novel compliant surgical devices. *Plos one*, 15(5):e0232952, 2020.

[219] Hao Liu, Amirhossein Farvardin, Robert Grupp, Ryan J Murphy, Russell H Taylor, Iulian Iordachita, and Mehran Armand. Shape tracking of a dexterous continuum manipulator utilizing two large deflection shape sensors. *IEEE sensors journal*, 15(10):5494–5503, 2015.

[220] Amirhossein Farvardin, Ryan J Murphy, Robert B Grupp, Iulian Iordachita, and Mehran Armand. Towards real-time shape sensing of continuum manipulators utilizing fiber bragg grating sensors. In *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, pages 1180–1185. IEEE, 2016.

[221] Xiaoguo Li, Lin Cao, Anthony Meng Huat Tiong, Phuoc Thien Phan, and Soo Jay Phee. Distal-end force prediction of tendon-sheath mechanisms for flexible endoscopic surgical robots using deep learning. *Mechanism and Machine Theory*, 134:323–337, 2019.

[222] Shahriar Sefati, Rachel Hegeman, Farshid Alambeigi, Iulian Iordachita, Peter Kazanzides, Harpal Khanuja, Russell H Taylor, and Mehran Armand. A surgical robotic system for treatment of pelvic osteolysis using an fbg-equipped continuum manipulator and flexible instruments. *IEEE/ASME Transactions on Mechatronics*, 26(1):369–380, 2020.

[223] Shahriar Sefati, Rachel Hegeman, Iulian Iordachita, Russell H Taylor, and Mehran Armand. A dexterous robotic system for autonomous debridement of osteolytic bone lesions in confined spaces: Human cadaver studies. *IEEE Transactions on Robotics*, 2021.

[224] Julie Legrand, Mouloud Ourak, Laura Van Gerven, Vincent Vander Poorten, and Emmanuel Vander Poorten. A miniature robotic steerable endoscope for maxillary sinus surgery called plient. *Scientific reports*, 12(1):1–15, 2022.

[225] Robert J Webster III and Bryan A Jones. Design and kinematic modeling of constant curvature continuum robots: A review. *The International Journal of Robotics Research*, 29(13):1661–1683, 2010.

[226] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

[227] Ta-Chih Lee, Rangasami L Kashyap, and Chong-Nam Chu. Building skeleton models via 3-d medial surface axis thinning algorithms. *CVGIP: Graphical Models and Image Processing*, 56(6):462–478, 1994.

[228] Philip HS Torr and Andrew Zisserman. Mlesac: A new robust estimator with application to estimating image geometry. *Computer vision and image understanding*, 78(1):138–156, 2000.

[229] Xuan Thao Ha, Mouloud Ourak, Omar Al-Ahmad, Di Wu, Gianni Borghesan, Arianna Menciassi, and Emmanuel Vander Poorten. Robust catheter tracking by fusing electromagnetic tracking, fiber bragg grating and sparse fluoroscopic images. *IEEE Sensors Journal*, 21(20):23422–23434, 2021.

[230] Thu Minh Le, Behzad Fatahi, Hadi Khabbaz, and Wenjing Sun. Numerical optimization applying trust-region reflective least squares algorithm with constraints to optimize the non-linear creep parameters of soft soil. *Applied Mathematical Modelling*, 41:236–256, 2017.

[231] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.

[232] Thomas F Banchoff and Stephen T Lovett. *Differential geometry of curves and surfaces*. CRC Press, 2016.

[233] Sungyeop Lim and Soonhung Han. Helical extension method for solving the natural equation of a space curve. *Surface Topography: Metrology and Properties*, 5(3):035002, 2017.

[234] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.

[235] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.

[236] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning internal representations by error propagation. Technical report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[237] Yi Zhou, Connelly Barnes, Jingwan Lu, Jimei Yang, and Hao Li. On the continuity of rotation representations in neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5745–5753, 2019.

[238] Afshin Gholamy, Vladik Kreinovich, and Olga Kosheleva. Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. 2018.

[239] Di Wu, Yao Zhang, Mouloud Ourak, Kenan Niu, Jenny Dankelman, and Emmanuel Vander Poorten. Hysteresis modeling of robotic catheters based on long short-term memory network for improved environment reconstruction. *IEEE Robotics and Automation Letters*, 6(2):2106–2113, 2021.

[240] Berthold KP Horn. Closed-form solution of absolute orientation using unit quaternions. *Josa a*, 4(4):629–642, 1987.

[241] Simon Baker, Takeo Kanade, et al. Shape-from-silhouette across time part i: Theory and algorithms. *International Journal of Computer Vision*, 62(3): 221–247, 2005.

[242] Yue Chen, Long Wang, Kevin Galloway, Isuru Godage, Nabil Simaan, and Eric Barth. Modal-based kinematics and contact detection of soft robots. *Soft Robotics*, 8(3):298–309, 2021.

[243] Isuru S Godage, Yue Chen, and Ian D Walker. Dynamic control of pneumatic muscle actuators. *IEEE IROS Workshop:Soft Robotic Modeling and Control*, 2018.

[244] Michaël De Volder, AJM Moers, and Dominiek Reynaerts. Fabrication and control of miniature mckibben actuators. *Sensors and Actuators A: Physical*, 166(1):111–116, 2011.

[245] Berengere Bardou, Florent Nageotte, Philippe Zanne, and Michel de Mathelin. Improvements in the control of a flexible endoscopic system. In *2012 IEEE International Conference on Robotics and Automation*, pages 3725–3732. IEEE, 2012.

[246] Olatunji Mumini Omisore, Shipeng Han, Tao Zhou, Yousef Al-Handarish, Wenjing Du, Kamen Ivanov, and Lei Wang. Learning-based parameter estimation for hysteresis modeling in robotic catheterization. In *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 5399–5402. IEEE, 2019.

[247] Rafael Aleluia Porto, Florent Nageotte, Philippe Zanne, and Michel de Mathelin. Position control of medical cable-driven flexible instruments by combining machine learning and kinematic analysis. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 7913–7919. IEEE, 2019.

[248] Paolo Cabras, Florent Nageotte, Philippe Zanne, and Christophe Doignon. An adaptive and fully automatic method for estimating the 3d position of bendable instruments using endoscopic images. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(4):e1812, 2017.

[249] Rob Reilink, Stefano Stramigioli, and Sarthak Misra. 3d position estimation of flexible instruments: marker-less and marker-based methods. *International journal of computer assisted radiology and surgery*, 8:407–417, 2013.

[250] TN Do, T Tjahjowidodo, MWS Lau, T Yamamoto, and SJ Phee. Hysteresis modeling and position control of tendon-sheath mechanism in flexible endoscopic systems. *Mechatronics*, 24(1):12–22, 2014.

[251] Olatunji Mumini Omisore, Shi Peng Han, Ling Xue Ren, Guo Shuai Wang, Fang Liang Ou, Hui Li, and Lei Wang. Towards characterization and adaptive compensation of backlash in a novel robotic catheter system for cardiovascular interventions. *IEEE Transactions on biomedical circuits and systems*, 12(4): 824–838, 2018.

[252] Julie Legrand, Dries Dirckx, Maarten Durt, Mouloud Ourak, Jan Deprest, Sebastien Ourselin, Jun Qian, Tom Vercauteren, and Emmanuel Vander Poorten. Active handheld flexible fetoscope–design and control based on a modified generalized prandtl-ishlinski model. In *2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 367–374. IEEE, 2020.

[253] Wenjun Xu, Jie Chen, Henry YK Lau, and Hongliang Ren. Data-driven methods towards learning the highly nonlinear inverse kinematics of tendon-driven surgical manipulators. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 13(3):e1774, 2017.

[254] Keunjun Choi, Jaewoon Kwon, Taeyoon Lee, Changwoo Park, Jinwon Pyo, Choongin Lee, SungPyo Lee, Inhyeok Kim, Sangok Seok, Yong-Jae Kim, et al. A hybrid dynamic model for the ambidex tendon-driven manipulator. *Mechatronics*, 69:102398, 2020.

[255] Noriaki Hirose and Ryosuke Tajima. Modeling of rolling friction by recurrent neural network using lstm. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6471–6478. IEEE, 2017.

[256] Vahid Hassani, Tegoeh Tjahjowidodo, and Thanh Nho Do. A survey on hysteresis modeling, identification and control. *Mechanical systems and signal processing*, 49(1-2):209–233, 2014.

[257] Saeid Shakiba, Mouloud Ourak, Emmanuel Vander Poorten, Moosa Ayati, and Aghil Yousefi-Koma. Modeling and compensation of asymmetric rate-dependent hysteresis of a miniature pneumatic artificial muscle-based catheter. *Mechanical Systems and Signal Processing*, 154:107532, 2021.

[258] Di Wu, Mouloud Ourak, Kenan Niu, Yao Zhang, M Ahmad, Jenny Dankelman, and Emmanuel Vander Poorten. Towards modeling of hysteresis in robotic catheters based on lstm. In *32nd Annual Conference of the International Society for Medical Innovation and Technology (iSMIT)*, pages 10–13, 2020.

[259] Shukuan Zhang, Mingqiao Wang, Ping Zheng, Guangyuan Qiao, Faliang Liu, and Lei Gan. An easy-to-implement hysteresis model identification method based on support vector regression. *IEEE Transactions on Magnetics*, 53 (11):1–4, 2017.

[260] Felix Bourier, Rebecca Fahrig, Paul Wang, Pasquale Santangeli, Klaus Kurzidim, Norbert Strobel, Teri Moore, Cameron Hinkel, and AMIN AL-AHMAD. Accuracy assessment of catheter guidance technology in electrophysiology procedures: a comparison of a new 3d-based fluoroscopy navigation system to current electroanatomic mapping systems. *Journal of cardiovascular electrophysiology*, 25(1):74–83, 2014.

[261] Jovana Janjic, Frits Mastik, Merel D Leistikow, Johan G Bosch, Geert Springeling, Antonius FW van der Steen, and Gijs van Soest. Sparse ultrasound image reconstruction from a shape-sensing single-element forward-looking catheter. *IEEE Transactions on Biomedical Engineering*, 65(10): 2210–2218, 2018.

[262] Jonas Smits, Mouloud Ourak, Andy Gijbels, Laure Esteveny, Gianni Borghesan, Laurent Schoevaerdts, Koen Willekens, Peter Stalmans, Eva Lankenau, Hinnerk Schulz-Hildebrandt, et al. Development and experimental validation of a combined fbg force and oct distance sensing needle for robot-assisted retinal vein cannulation. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 129–134. IEEE, 2018.

[263] Francesco Sturla, Riccardo Vismara, Michal Jaworek, Emiliano Votta, Paolo Romitelli, Omar A Pappalardo, Federico Lucherini, Carlo Antona, Gianfranco B Fiore, and Alberto Redaelli. In vitro and in silico approaches to quantify the effects of the mitraclip® system on mitral valve function. *Journal of biomechanics*, 50:83–92, 2017.

[264] Siddharth Krishna Kumar. On weight initialization in deep neural networks. *arXiv preprint arXiv:1704.08863*, 2017.

[265] Samuel B Kesner and Robert D Howe. Force control of flexible catheter robots for beating heart surgery. In *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1589–1594, 2011.

[266] Mohsen Mahvash and Pierre E Dupont. Stiffness control of surgical continuum manipulators. *IEEE Trans. on Robot.*, 27(2):334–345, 2011.

[267] Roger E Goldman, Andrea Bajo, and Nabil Simaan. Compliant motion control for multisegment continuum robots with actuation force sensing. *IEEE Transactions on Robotics*, 30(4):890–902, 2014.

[268] Guanglin Ji, Junyan Yan, Jingxin Du, Wanquan Yan, Jibiao Chen, Yongkang Lu, Juan Rojas, and Shing Shin Cheng. Towards safe control of continuum manipulator using shielded multiagent reinforcement learning. *IEEE Robotics and Automation Letters*, 6(4):7461–7468, 2021.

[269] Michael C Yip and David Camarillo. Model-less feedback control of continuum manipulators in constrained environments. *IEEE Transactions on Robotics*, 30(4):880–889, 2014.

[270] Michael C Yip et al. Model-less hybrid position/force control: a minimalist approach for continuum manipulators in unknown, constrained environments. *IEEE Robot. and Automat. Lett.*, 1(2):844–851, 2016.

[271] David Jakes, Zongyuan Ge, and Liao Wu. Model-less active compliance for continuum robots using recurrent neural networks. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

[272] Milad Azizkhani, Isuru S Godage, and Yue Chen. Dynamic control of soft robotic arm: A simulation study. *IEEE Robot. and Automat. Lett.*, 7(2): 3584–3591, 2022.

[273] Berengere Bardou, Florent Nageotte, Philippe Zanne, and Michel de Mathelin. Improvements in the control of a flexible endoscopic system. In *2012 IEEE International Conference on Robotics and Automation*, pages 3725–3732. IEEE, 2012.

[274] Bingbin Yu, José de Gea Fernández, and Tao Tan. Probabilistic kinematic model of a robotic catheter for 3d position control. *Soft robotics*, 6(2):184–194, 2019.

[275] Donghoon Baek, Ju-Hwan Seo, Joonhwan Kim, and Dong-Soo Kwon. Hysteresis compensator with learning-based hybrid joint angle estimation for flexible surgery robots. *IEEE Robotics and Automation Letters*, 5(4): 6837–6844, 2020.

[276] TN Do, T Tjahjowidodo, MWS Lau, T Yamamoto, and SJ Phee. Hysteresis modeling and position control of tendon-sheath mechanism in flexible endoscopic systems. *Mechatronics*, 24(1):12–22, 2014.

[277] Xiangyu Wang, Dongyang Bie, Jianda Han, and Yongchun Fang. Active modeling and compensation for the hysteresis of a robotic flexible ureteroscopy. *IEEE Access*, 8:100620–100630, 2020.

[278] Saeid Shakiba, Aghil Yousefi-Koma, and Moosa Ayati. Tracking control of an sma-driven actuator with rate-dependent behavior using an inverse model of hysteresis. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 42:1–15, 2020.

[279] Bryan A Jones and Ian D Walker. Kinematics for multisection continuum robots. *IEEE Transactions on Robotics*, 22(1):43–55, 2006.

[280] Vahid Hassani, Tegoeh Tjahjowidodo, and Thanh Nho Do. A survey on hysteresis modeling, identification and control. *Mech. syst. and signal process.*, 49(1-2):209–233, 2014.

[281] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[282] H Nijland et al. Evaluation of accuracy and precision of ct-guidance in radiofrequency ablation for osteoid osteoma in 86 patients. *PLoS One*, 12(4): e0169171, 2017.

[283] Roger E Goldman, Andrea Bajo, and Nabil Simaan. Compliant motion control for continuum robots with intrinsic actuation sensing. In *2011 IEEE Int. Conf. on Robot. and Automa. (ICRA)*, pages 1126–1132. IEEE, 2011.

[284] Herman Bruyninckx. Open robot control software: the orocos project. In *Proceedings 2001 ICRA. IEEE international conference on robotics and automation (Cat. No. 01CH37164)*, volume 3, pages 2523–2528. IEEE, 2001.

[285] Mohammad Hasan Dad Ansari et al. Proof-of-concept medical robotic platform for endovascular catheterization. In *11th Conf. on New Technologies for Computer and Robot Assisted Surgery (CRAS)*, pages 66–67, 2022.

[286] Xiaobo Tan and Ram V Iyer. Modeling and control of hysteresis. *IEEE Control Systems Magazine*, 29(1):26–28, 2009.

[287] Fabian Trauzettel, Mohammad Hasan Dad Ansari, Beatriz Farola Barata, Di Wu, and Zhen Li. Exploitation plan 4 of atlas project deliverable d7.11, 2022.

[288] Catheters market size, share and trends | global industry growth 2025, . URL `https://www.alliedmarketresearch.com/catheters-market`.

[289] Catheter market size & share | industry growth report, 2019-2026, . URL `https://www.grandviewresearch.com/industry-analysis/catheters-market-analysis`.

[290] Catheters market size, share poised to touch usd 56.5 billion by 2025 - 12 nov 2019, . URL `https://www.marketwatch.com/press-release/catheters-market-size-share-poised-to-touch-usd-565-billion-by-2025---12-nov-2019-2019-11-11`.

[291] Catheters market - forecast to 2025 | by type & end user, . URL `https://www.marketsandmarkets.com/Market-Reports/catheters-market-6247803.html`.

[292] Catheters market research forecasts 2025 | regional outlook, . URL `https://www.gminsights.com/industry-analysis/catheters-market`.

[293] Song-kyoo Kim. Design of enhanced software protection architecture by using theory of inventive problem solving. In *2009 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 978–982. IEEE, 2009.

[294] Sergiy Gnatyuk, Vasyl Kinzeryavyy, Iryna Stepanenko, Yana Gorbatyuk, Andrii Gizun, and Vitalii Kotelianets. Code obfuscation technique for enhancing software protection against reverse engineering. In *Advances in Artificial Systems for Medicine and Education II 2*, pages 571–580. Springer, 2020.

[295] Christian S. Collberg and Clark Thomborson. Watermarking, tamper-proofing, and obfuscation-tools for software protection. *IEEE Transactions on software engineering*, 28(8):735–746, 2002.

[296] Gholamreza Afsharipour. The nature and characteristics of the license agreement and its compatibility with certain contracts in civil law. *Mediterranean Journal of Social Sciences*, 7(4 S1):105, 2016.

[297] Sofia Palmieri, Paulien Walraet, and Tom Goffin. Inevitable influences: Ai-based medical devices at the intersection of medical devices regulation and the proposal for ai regulation. *European Journal of Health Law*, 28(4): 341–358, 2021.

[298] Feiyu Xu, Hans Uszkoreit, Yangzhou Du, Wei Fan, Dongyan Zhao, and Jun Zhu. Explainable ai: A brief survey on history, research areas, approaches and challenges. In *Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, Proceedings, Part II 8*, pages 563–574. Springer, 2019.

[299] Parvathaneni Naga Srinivasu, N Sandhya, Rutvij H Jhaveri, and Roshani Raut. From blackbox to explainable ai in healthcare: existing tools and case studies. *Mobile Information Systems*, 2022:1–20, 2022.

[300] Di Wu, Yao Zhang, Mouloud Ourak, Xuan Thao Ha, Kenan Niu, Jenny Dankelman, and Emmanuel Vander Poorten. Deep-learning-based position control of a robotic catheter under environmental contact. In *2022 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE, 2022.

[301] Jorge F Lazo, Chun-Feng Lait, Sara Moccia, Benoit Rosa, Michele Catellani, Michel de Mathelin, Giancarlo Ferrigno, Paul Breedveld, Jenny Dankelman, and Elena De Momi. Autonomous intraluminal navigation of a soft robot using deep-learning-based visual servoing. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6952–6959. IEEE, 2022.

[302] C Lai, Fabian Trauzettel, Paul Breedveld, Elena De Momi, Giancarlo Ferrigno, et al. Modularization of medical robotic manipulator for adapting soft robotic arms with varying numbers of dofs. In *Proc. of the 10th Conference on New Technologies for Computer and Robot Assisted Surgery (CRAS), 2020*, 2020.

[303] Christoff M Heunis, Kasper J Behrendt, Edsko EG Hekman, Cyril Moers, Jean-Paul PM de Vries, and Sarthak Misra. Design and evaluation of a magnetic rotablation catheter for arterial stenosis. *IEEE/ASME transactions on mechatronics*, 27(3):1761–1772, 2021.

[304] Giovanni Pittiglio, Peter Lloyd, Tomas da Veiga, Onaizah Onaizah, Cecilia Pompili, James H Chandler, and Pietro Valdastri. Patient-specific magnetic catheters for atraumatic autonomous endoscopy. *Soft Robotics*, 9(6):1120–1133, 2022.

[305] Chayabhan Limpabandhu, Yihua Hu, Hongliang Ren, Wenzhan Song, and Zion Tsz Ho Tse. Magnetically steerable catheters: State of the art review. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 237(3):297–308, 2023.

[306] D Usvald, J Hlucilova, J Strnadel, R Prochazka, J Motlik, M Marsala, et al. Permanent jugular catheterization in miniature pig: treatment, clinical and pathological observations. *Veterinární medicína*, 53(7):365–372, 2008.

Acknowledgements

I finally started writing the acknowledgment of my PhD thesis. Reflecting on my PhD journey, I realize that it has not been as difficult as I initially thought. By taking a step-by-step approach, I am convinced that anyone can successfully complete their PhD journey. My perseverance, I feel, is largely due to the support of my family, supervisors, and friends, and the nurturing environment they have provided. From what I understand, acknowledgment is often a place to express thanks, but in my thesis, I want to deviate slightly from this tradition. I firmly believe that gratitude towards someone should be expressed in everyday life, not just vocally acknowledged in a document. So I plan to add a twist to my acknowledgment, first sharing my current thoughts and reflections.

Surgical robotics, a fascinating topic that always attracts the admiration of my friends when discussing my work. They perceive me as someone who is engaged in high-end and sophisticated technology. Of course, they might imagine that I am creating humanoid robots for surgery, like in sci-fi movies. I do feel a sense of pride working in this field. Growing up, I was not particularly drawn to medicine, although my father was a radiologist and our home was filled with medical books. My parents never really pushed me towards a medical career, thinking it was too demanding and exhausting. So I chose mechanical engineering in my undergraduate studies, thinking that it would take me away from medicine. But fate had other plans, and I ended up specializing in medical technology during my Master's program. This led to a shift towards surgical robotics, a field I continued into my PhD and likely beyond. I am quite happy with this path. Now, when I talk with my father, our conversations are richer, often centered on the medical devices I am working on. He likes to offer his insights and suggest research directions, and I treasure these interactions. It has been eight years since I started studying abroad and being far from home, I feel that I have missed many opportunities to connect with my family. I hope to spend more time with them in the coming years.

Finding time to spend with family is not always easy, and *I have observed in myself a growing tendency to become a workaholic*, although I try to curb this tendency. I vividly recall how thrilled I was when my first paper was accepted for ISMR while I was in the US. The joy of receiving that acceptance email stayed with me for weeks. In contrast, I remained very calm upon receiving the acceptance of my RAL and TRO papers. This experience made me realize that *people tend to become greedy.* Remembering where we started and learning to be content is the key to a happy and meaningful life. During my PhD, I have been fortunate to attend various conferences such as ICRA, IROS, CRAS, and the Hamlyn Symposium. I have had the chance to speak with many professors, who often seem in a hurry. A prime example was at IROS 2022 in Japan, where professors traveled for hours from Europe or the US. Despite the long trip, they opted to stay in Japan for only few days. This seems typical in academia, so learning work-life balance might be the most crucial lesson for people working in this field. I often find myself reflecting on what I am truly striving for? freedom, in particular, *the freedom of time.* This kind of freedom, to me, holds far greater value than the fleeting satisfaction of publishing papers or earning promotions. I see some people relentlessly working, without any time to relax or travel. That is not the life I envision for myself. I hope to be driven by inspiration in my work and also cherish the freedom to take breaks or travel at my own leisure. Therefore, in my future life, I hope to learn how to temper my ambitions, and engage more in activities that bring me joy, and live in harmony with my true self.

There is a saying that I particularly like: *"Life is like a book, and those who do not travel only read one page of it."* I have a passion for traveling and exploring different countries. I was born and raised in a small mining town where I spent the most carefree days of my life. Although urbanization has thinned its population, I often find myself returning there, feeling a genuine sense of belonging to that place. After completing my undergraduate studies in China, I moved to Germany for my Master's degree. During this time, I spent a semester at EPFL in Lausanne, Switzerland. At the tail end of my master's study, I went to Johns Hopkins University in the US for my master thesis. Later, I returned to Europe to start my double PhD program in Belgium and the Netherlands. During my doctoral studies, I went to Italy for a short exchange due to project collaboration. My academic journey has been exciting, and traveling through these countries broadened my horizons. I gained a deep understanding of the local customs, culture, and character of the residents in each place, which you cannot grasp through social media. I hope that in the future, I can carry on with this adventure and explore many more countries.

After penning my message to my future self, it is time to express gratitude to others who have assisted me on this journey. I wish to extend my appreciation to my supervisor, Emmanuel Vander Poorten. The first time I met him was at ISMR 2019 in the US. Prior to the conference, I reached out to Manu via email, hoping for a brief meeting during the event. Our conversation there was short but impactful. At ISMR 2019, I delivered my first scientific presentation in English, after which Manu, as an audience member, asked the first question. His continued engagement was apparent when he sent me several follow-up questions via email post-presentation, showcasing his passion for research. I am profoundly grateful to Manu for offering me the PhD opportunity. His constructive feedback over the

four years was invaluable and significantly contributed to my work. His diligent work ethic has deeply inspired and influenced me.

I would like to extend my heartfelt appreciation to another supervisor, Jenny Dankelman. She excels not only as an outstanding supervisor but also shines as a person of genuine kindness and warmth. From our initial meeting, Jenny never felt like a typical full professor to me, but rather a friend. Throughout the past four years, her guidance has gone beyond just scientific feedback. She has also offered invaluable life advice, such as maintaining a work-life balance. Her enthusiasm for tutoring and interacting with students is evident, and she has consistently encouraged me to reflect and strive for a balanced life. What struck me most was her attention to detail regarding my defense preparations. She often seemed more concerned than I was about the procedure, proactively emailing me reminders and guidance for the next steps. I am deeply grateful to have had such a wonderful supervisor in my academic journey.

I would also like to express my gratitude to Iulian Iordachita, who supervised my master's thesis at Johns Hopkins University in the USA. I vividly recall the days spent in the lab, sitting in front of my computer, while Iulian, dressed in his blue work uniform, diligently worked on assembly tasks behind me. His hands-on approach deeply influenced me. I still remember the times when Iulian would sit beside me, asking me to open Overleaf, and together, line by line, we spent three hours to refine the first paper I ever wrote. I am deeply grateful to Iulian for setting a remakable example and for providing me with an entry onto a world-class stage for research, opening the doors to the expansive world of academia. Thank you, Iulian, for your invaluable mentorship and guidance!

Thanks to Professor Arianna Menciassi from SSSA for giving me the opportunity to visit her lab. I was fortunate to attend her course “fundamentals of surgical robots”, which discussed the impressive history of surgical robot development and was incredibly inspiring. I would also like to extend my thanks to Paul Breedveld, who is much more than a professor – he is a talented photographer, and someone always willing to lend a helping hand. I distinctly remember the valuable lessons he taught me on image presentation and the stunning photographs he created for his group website.

A big thank you to all the professors on my defense committee: Pádraig Cantillon-Murphy, Aimée Sakes, Vero Vanden Abeele, Jos Vander Sloten, Helge Wurdemann, Jules Scheltes, and Gianni Borghesan. Their dedication in meticulously reviewing my thesis and providing insightful feedback significantly enriched the quality of my work. Additionally, I am grateful for their time and effort in attending and contributing to my defense, making it an enlightening experience.

Of course, I would like to thank Mouloud, who closely monitored my progress throughout my PhD journey and offered lots of support. I am also grateful to all members of the RAS group, with a special mention to my Chinese friends. Our mutual support was a cornerstone in alleviating much of the stress associated

with the PhD process. Heartfelt thanks to you all! Additionally, I extend my appreciation to the members of the TU Delft MISIT group and the Surgical Robotics and Allied Technologies group at SSSA for their invaluable assistance and hospitality during my visiting period.

I am deeply indebted to the EU Marie Skłodowska-Curie ATLAS project, which played a pivotal role during my PhD journey. As a Marie Curie ITN fellow, I was given the extraordinary opportunity to pursue dual PhD degrees. This journey also brought me into collaboration with many brilliant minds in the field, enriching my experience significantly. Without the ATLAS project, my PhD path would undoubtedly have been less dynamic and engaging.

Above all, my deepest gratitude goes to my Mama, Papa, and my beloved. Thank you for your unwavering support throughout my PhD journey. Though I strive to express my gratitude in our everyday moments, words often fall short of capturing the full depth of my feelings.

My PhD journey is nearing its conclusion. After many years, I hope to look back with satisfaction on these enriching experiences of studying and traveling around the globe. *I am grateful that, during my PhD studies, I chose not to overwhelm myself with pressure. Instead, I allowed ample time for relaxation, travel, and experiencing different countries, rather than staying confined to the lab or home and becoming overly addicted to the thrill of publishing papers.*

As I turn the final page of this PhD journey, I step into the next phase of my life, armed with knowledge and experience. I am fully aware that numerous challenges and joys await me. Guided by the principle, *"Be happy with what you have while working for what you want,"* I am prepared to face life's joys and challenges with optimism and determination. Years from now, I hope to reflect on this academic odyssey, which spanned multiple countries and was filled with joy, challenges, and invaluable learning, with a profound sense of pride.

WU Di
Leuven
November 2023

Curriculum vitae

Di Wu was born in Anhui, China. He received the B.Sc. degree from Wuhan University of Technology, China. After this, he received the M.Sc. degree in medical technology and engineering from the Technical University of Munich (TUM), Germany. He is currently pursuing a double Ph.D. degree at KU Leuven, Belgium, and Delft University of Technology, the Netherlands, under the Marie Skłodowska-Curie Actions Program of the European Union Horizon 2020.

From 2017 to 2018, he was an exchange student at the École Polytechnique Fédérale de Lausanne (EPFL), funded by the Swiss European Mobility Programme (SEMP). From 2018 to 2019, he was a Visiting Graduate Researcher at the Laboratory for Computational Sensing and Robotics (LCSR) at Johns Hopkins University, USA. In 2023, he was a visiting Ph.D. student at The BioRobotics Institute of Scuola Superiore Sant'Anna, Pisa, Italy.

His paper received an Honorable Mention for the Best Paper Award from IEEE Robotics and Automation Letters (RA-L) in 2021 (10 out of more than 1100 papers), and another was a finalist for the Best Paper Award at the IEEE International Symposium on Medical Robotics (ISMR) in 2019.

He co-organized two workshops: one at the 2022 IEEE/RSJ International Conference on Intelligent Robots (IROS) and another at the 2023 Hamlyn Symposium on Medical Robotics.

List of publications

Articles in internationally reviewed academic journals

- **Wu, D.** , Li, Z. , Ansari, M. H. D., Ha, X. T. , Ourak, M. , Dankelman, J. , Menciassi, A. , De Momi, E. , Vander Poorten, E. , with Wu, D. (joint first author), Li. Z (joint first author), (2023). Comparative Analysis of Interactive Modalities for Intuitive Endovascular Interventions. (submitted)
- **Wu, D.**, Zhang, R., Pore, A., Dall'Alba, D., Ha, X. T., Li, Z., Zhang, Y., Gonzalez, F., Ourak, M., Kowalczyk, W., De Momi, E., Casals, A., Dankelman, J., Kober, J., Menciassi, A., Fiorini, P., Vander Poorten, E., with Wu, D. (joint first author), Zhang. R (joint first author), (2023). A Review on Machine Learning in Flexible Surgical and Interventional Robots: Where We Are and Where We Are Going. (submitted)
- **Wu, D.**, Ha, X.T., Zhang, Y., Ourak, M., Borghesan, G., Niu, K., Trauzettel, F., Dankelman, J., Menciassi, A., Vander Poorten, E. with Wu, D. (joint first author), Ha, X.T. (joint first author) (2022). Deep-learning-based Compliant Motion Control of a Pneumatically-driven Robotic Catheter. *IEEE Robotics and Automation Letters,* 1-8. doi: 10.1109/LRA.2022.3186497
- Ha, X.T., **Wu, D.**, Ourak, M., Borghesan, G., Dankelman, J., Menciassi, A., Vander Poorten, E. with Ha, X.T. (joint first author), Wu, D. (joint first author) (2022). Shape Sensing of Flexible Robots based on Deep Learning. *IEEE Transactions On Robotics.* doi: 10.1109/TRO.2022.3221368
- **Wu, D.**, Zhang, Y., Ourak, M., Niu, K., Dankelman, J., Vander Poorten, E. with Wu, D. (joint first author), zhang, Y. (joint first author) (2021). Hysteresis Modeling of Robotic Catheters based on Long Short-Term Memory Network for Improved Environment Reconstruction. *IEEE Robotics and Automation Letters,* 6 (2), 2106-2113. doi: 10.1109/LRA.2021.3061069 **(Honorable Mentions of the Best Paper Award of IEEE RA-L in year 2021)**
- Zhang, Y., Verschooten, E., Ourak, M., Van Assche, K., Borghesan, G., **Wu, D.**, Niu, K., Joris, P.X. and Vander Poorten, E., (2023). Physiological Motion

Compensation for Neuroscience Research based on Electrical Bio-Impedance Sensing. *IEEE Sensors Journal.*

- Ha, X.T., **Wu, D.**, Ourak, M., Borghesan, G., Menciassi, A., Vander Poorten, E. (2023). Sensor Fusion for Shape Reconstruction using Electromagnetic Tracking Sensors and Multi-core Optical Fiber. *IEEE Robotics and Automation Letters.* doi: 10.1109/LRA.2023.3280456

- Ha, X.T., **Wu, D.**, Lai, C-F., Ourak, M., Borghesan, G., Menciassi, A., Vander Poorten, E. (2022). Contact Localization of Continuum and Flexible Robot Using Data-driven Approach. *IEEE Robotics and Automation Letters.* doi: 10.1109/LRA.2022.3176723

- Ha, X.T., Ourak, M., Al-Ahmad, O., **Wu, D.**, Borghesan, G., Menciassi, A., Vander Poorten, E. (2021). Robust Catheter Tracking by Fusing Electromagnetic Tracking, Fiber Bragg Grating and Sparse Fluoroscopic Images. *IEEE Sensors Journal,* 1-12. doi: 10.1109/JSEN.2021.3107036

Papers disseminated at international scientific conferences and symposia, published in full in proceedings

- **Wu, D.**, Sridhar, A., Beckers, W-A., Mehdi, S. Z., Ourak, M., Tamadon, I., Dankelman, J., Votta, E., Menciassi, A., Vander Poorten, E., (2023). "Data-driven Modeling of Complex Hysteresis Behaviour in MitraClip Steerable Catheters", *12th Conference on New Technologies for Computer and Robot Assisted Surgery.*

- **Wu, D.**, Zhang, Y., Ourak, M., Xuan, T.H., Niu, K., Dankelman, J., Vander Poorten, E. with Wu, D. (joint first author), zhang, Y. (joint first author) (2022). Deep-learning-based Position Control of a Robotic Catheter under Environmental Contact. *2022 International Symposium on Medical Robotics (ISMR),* GA, Atlanta, 13 Apr 2022-15 Apr 2022. doi: 10.1109/ISMR48347.2022.9807533

- **Wu, D.**, Ourak, M., Niu, K., Zhang, Y., Ahmad, M.A., Dankelman, J., Vander Poorten, E. (2020). Towards Modeling of Hysteresis in Robotic Catheters based on LSTM. *32nd Conference of the International Society for Medical Innovation and Technology (iSMIT),* Chicago, USA, 03 Dec 2020-10 Dec 2020.

- **Wu, D.**, Ourak, M., Ahmad, M.A., Niu, K., Borghesan, G., Dankelman, J., Vander Poorten, E. (2020). Feasibility of using a Long Short-Term Memory Network for Robotic Catheter Control. (68-69). *10 th Conference on New Technologies for Computer and Robot Assisted Surgery,* Barcelona, Spain, 28 Sep 2020-30 Sep 2020.

- Trauzettel, F., De Koning, A., **Wu, D.**, Vander Poorten, E., Breedveld, P., (2023). Creating Soft Robotic Manipulators with Variable Stiffness via 3D Printing and Fiber Jamming. *12th Conference on New Technologies for Computer and Robot Assisted Surgery.*
- Mohammad, H.D A., Farola Barata, B., Trauzettel, F., Li, Z., **Wu, D.**, Dall'Alba, D., Borghesan, G., Ourak, M., Iacovacci, V., Tognarelli, S., Dankelman, J., De Momi, E., Breedveld, P., Fiorini, P., Vander Sloten, J., Menciassi, A., Vander Poorten, E. (2022). Proof-of-Concept Medical Robotic Platform for Endovascular Catheterization. (66-67). *2022 Conference on New Technologies for Computer/Robot Assisted Surgery,* Napoli, Italy.
- Ha, X.T., **Wu, D.**, Ourak, M., Borghesan, G., Menciassi, A., Vander Poorten, E. (2022). Compliant Motion Control of Robotic Catheter based on Long-Short Term Memory Network. *11th Conference on New Technologies for Computer/Robot Assisted Surgery* . Naples, Italy, 25 Apr 2022-27 Apr 2022.
- Zhang, Y., Ourak, M., Verschooten, E., Niu, K., **Wu, D.**, Borghesan, G., Joris, P., Vander Poorten, E. (2022). Controlling of a Micropositioning Piezoelectric Actuator using an LSTM Network for Robotic Eye Surgery. *2022 Conference on New Technologies for Computer and Robot Assisted Surgery*, Napoli, Apr. 2022
- Li, R., Niu, K., **Wu, D.**, Vander Poorten, E. (2020). A Framework of Real-time Freehand Ultrasound Reconstruction based on Deep Learning for Spine Surgery. *10 th Conference on New Technologies for Computer and Robot Assisted Surgery,* Barcelona, Spain, 28 Sep 2020-30 Sep 2020.

Propositions

accompanying the dissertation

CONTROL OF MULTI-DEGREE-OF-FREEDOM CATHETERS IN UNKNOWN ENVIRONMENTS

EXPLORING THE POTENTIAL OF DEEP LEARNING AND AUGMENTED REALITY

by

Di Wu

1. Within the next 10 years, we will witness the advent of autonomous surgery. [This thesis]
2. Public understanding of robotic surgery is very limited. It is urgent for us researchers to engage more in scientific popularization. [This thesis]
3. Efforts invested in deep learning methods are consistently undervalued by peers in the field.
4. As a PhD student, the ultimate measure of your work lies in its potential for product translation, rather than publishing in top-tier journals.
5. When we dislike something, such as deep learning, it is more insightful to examine its downsides in depth, instead of disregarding it superficially.
6. The policy of sequential listing for co-first authors unfairly disadvantages the secondly listed one. This practice is not good for collaboration and requires improvement.
7. Choosing industry over academia is a more effective path to turn ideas into reality.
8. Life's greatest lesson is learning to be satisfied.
9. *"The world is a book, and those who do not travel read only one page."* — Saint Augustine
10. Learning to prioritize work-life balance should be the first lesson for everyone in academia.
11. A dual PhD degree teaches important skills, especially negotiation between two parties.

These propositions are regarded as opposable and defendable, and have been approved as such by the promotors prof. dr. J. Dankelman, and dr. ir. E. Vander Poorten.