Title	情報駆動型触覚センサーシステムの開発
Author(s)	長澤, 知史
Citation	
Issue Date	2026-03
Type	Thesis or Dissertation
Text version	ETD
URL	https://hdl.handle.net/10119/20591
Rights	
Description	Supervisor: HO, Anh Van, 先端科学技術研究科, 博士

Development of Data-driven Vision-Based Tactile Sensor System

by

Satoshi Nagasawa

A Doctor Thesis

Graduate School of Advanced Science and Technology

Japan Advanced Institute of Science and Technology

(Information Science)

Thesis Supervisor: Ho Anh Van

March ,2026

ABSTRACT

Minimally invasive surgeries are performed using various medical instruments such as catheters and endoscopes in order to reduce the burden on patients. However, these procedures rely on indirect observation modalities, including endoscopy and X-ray CT, and therefore cannot be performed in an intuitively manipulable manner. As a result, a high level of operator expertise is required. In this context, the introduction of tactile sensing into minimally invasive medical devices is expected to improve the spatial resolution of intraoperative monitoring.

To date, a number of catheter-type tactile sensors based on electrical principles have been proposed, but all of them face challenges when high-density tactile sensing is required. In this study, we address these challenges by proposing a novel vision-based tactile sensing system, “TacBalloon” , which is built upon a balloon catheter originally developed for cardiac ablation procedures.

In particular, to realize a compact, balloon-catheter–scale tactile sensor, which has been difficult to achieve using conventional vision-based tactile sensing methodologies, we construct a data-driven tactile sensing framework. The proposed system adopts a marker-based tactile sensing approach, in which a marker pattern printed on the skin surface is observed through an endoscope and used to infer tactile information. Tactile estimation is performed by a neural-network-based tactile estimation model, which is trained on a dataset generated entirely in simulation. During training, data augmentation and related techniques are introduced to mitigate the Sim2Real gap.

We experimentally evaluated the performance of the tactile balloon catheter implemented using the proposed data-driven tactile sensing system. Within regions where the surface markers were visible, the system achieved a contact depth estimation accuracy of approximately ±1 mm and reliably reconstructed complex surface deformations. Furthermore, by comparing against photogrammetry, we assessed its ability to estimate contact areas on anatomically realistic models. The results demonstrate that TacBalloon enables high-density, wide-area tactile sensing on a flexible balloon surface and constitutes a promising solution for providing real-time tactile feedback in minimally invasive surgery.

Keyword:Tactile sensing, Tactile sensor, Vision-Based Tactile sensor, Balloon Catheter, Softbody Simulation, Optical Simulation , Machine learning

Acknowledgements

I would like to express my sincere gratitude to Associate Professor Ho Anh Van of the School of Materials Science, Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology, for his warm guidance and numerous valuable insights throughout the course of this research. His advice has been invaluable in shaping both the direction of the study and my academic perspective.

I also wish to extend my heartfelt thanks to all collaborators at Toray Industries, Inc. , including the members of the Cardiac Ablation Catheter Development Office, who have greatly supported this work as part of a joint research project. Their provision of experimental environments, technical expertise, and continual cooperation were essential to the progress of this study.

This research was supported by the Japan Science and Technology Agency’ s Support for Pioneering Research Initiated by the Next Generation (SPRING) program (JPMJSP2102). I gratefully acknowledge this support.

I am additionally indebted to the members of my laboratory for their continuous encouragement and for the insightful discussions that enriched my research activities on a daily basis.

Finally, I would like to express my deepest appreciation to my parents, whose understanding and unwavering support made it possible for me to pursue doctoral studies.

Table of Contents

Chapter3 Data-driven tactile balloon catheter 37

Chapter4 Experiments and Performance Evaluation 64

Table of Figures

Table of Tables

Chapter1

Introduction

1.1 Background

1.1.1 Tactile Sensing

In recent years, advances in the field of robotics have expanded the range of tasks that robots can perform from conventional manufacturing to logistics, agriculture, service industries, and even highly specialized domains such as medical care and nursing care. Against the backdrop of global labor shortages and population aging, demand for automation and labor-saving through the use of robots has been steadily increasing. According to a 2010 report by the New Energy and Industrial Technology Development Organization (NEDO), a national research and development agency in Japan, the robotics market, including its latent potential, was projected to reach 5.3 trillion JPY (60.4 billion USD) by 2025 and 9.7 trillion JPY (110.9 billion USD) by 2035, based on the average 2010 exchange rate of 87.8 JPY per USD [1]. Estimates by private market research firms for around 2025 suggest that, even when limited to industrial robots, the market size of the robotics industry will be on the order of 50 billion USD in 2025, and is expected to grow to 90.6-–111.0 billion USD by 2030 and 291.1 billion USD by 2035 [2–4]. The market for service robots is also projected to surpass NEDO’ s earlier forecasts [5], indicating an overall acceleration in market growth . NEDO’ s 2024 market analysis identifies the manufacturing sector, warehouse management and logistics, and the medical, nursing, and welfare sectors as the main segments driving the expansion of the AI robotics market, suggesting that technological innovation in these areas will be a key factor in transforming future social infrastructure [6].

Table 1.1: Market estimates and projections for the overall robotics industry (billion USD). Values for 2025 from NEDO (2010) are projections, while those from private research firms are estimates.

Source	2025	2030	2035	Segment	Reference
NEDO (2010)	60.4	-	110.9	Total	[1]
	29.8	-	54.1	Industrial[1]	[1]
	30.7	-	56.8	Service	[1]
ABI Research (2025)	44.8	110.7	-	Industrial	[2]
Mordor Intelligence (2025)	48.3	90.6	-	Industrial	[4]
Future Market Insights (2025)	55.1	-	291.1	Industrial	[3]
	58.5	-	393.3	Service	[5]

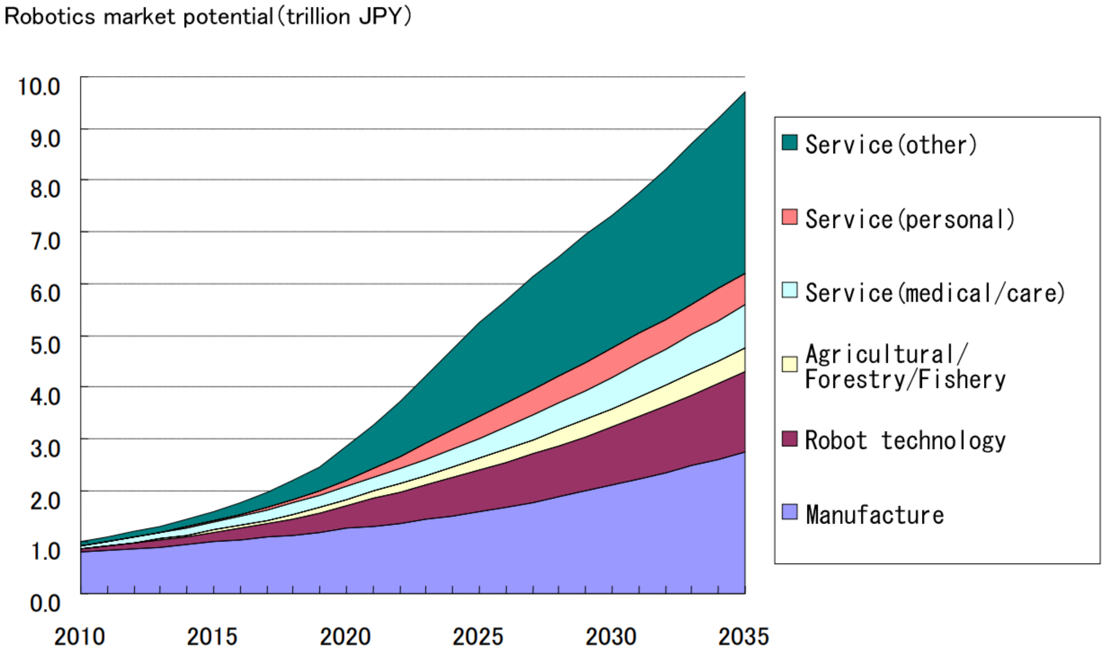


Fig. 1.1: Robotics market forecast reported by NEDO (2010–2035) [1]. English translation by the author. : The service sector is projected to exhibit particularly strong growth.

Against this background, one of the challenges that has become apparent in the deployment of industrial robots and medical/care robots is the assurance of safety and environmental adaptability. In particular, for collaborative robots that are designed to work in proximity with humans, and for medical and nursing-care robots that interact with the interior of the human body, it is technically crucial to reduce the risk that the robot may damage humans or the surrounding environment. Moreover, when handling objects that are soft and easily crumbled, such as food products, deformable flexible materials

---

[1] Within the NEDO forecast, the industrial segment is obtained by summing the manufacturing, robot technology, and non-service agricultural/forestry/fishery sectors reported in [1].

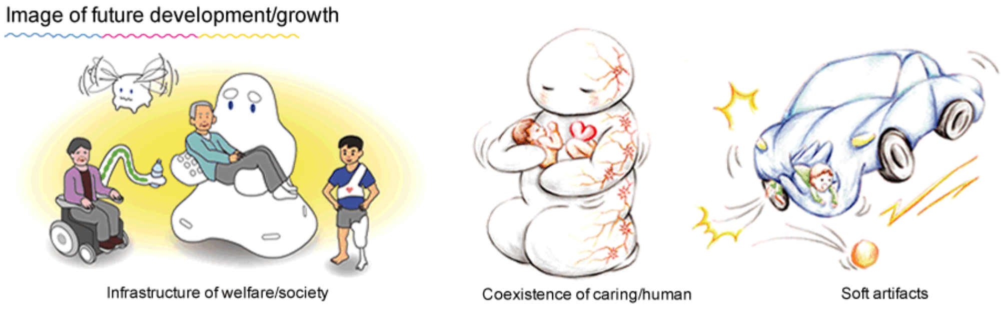


Fig. 1.2: Image of future soft robot [7]

such as rubber and fabric, or granular media that are inherently difficult to grasp, conventional robots based on rigid links and motor-driven actuation face significant limitations in adaptability [6]. Motivated by these issues, soft robots that actively incorporate flexible materials into their structures and actuation mechanisms have attracted considerable attention. Leveraging their intrinsic compliance, soft robots enable enhanced safety and novel approaches to robotic manipulation that exploit passive deformation to achieve better conformity to target objects.

Alongside the progress in soft robotics, the importance of tactile sensing—which allows robots to physically interact with the external world and “perceive” contact states—has been rapidly increasing. Tactile sensing refers to technologies that measure mechanical interactions between the sensor and the object (e.g., normal forces, shear forces, pressure distributions, shape deformation, and contact locations) and feed this information back into robot control and inference. Analogous to the tactile information obtainable from human skin on the fingers, tactile sensing is expected to provide robots with direct information about “what is touching where and how,” thereby improving the accuracy and safety of various tasks such as grasping, exploration, and manipulation.

A wide range of approaches has been proposed to realize tactile sensing, including electrical, acoustic, fluidic, and optical methods. Electrical approaches utilize piezoresistive, capacitive, and piezoelectric effects to convert external forces into electrical signals, enabling highly sensitive and high-speed sensing. In combination with MEMS technologies, these approaches have enabled high-density, miniaturized sensor arrays [8, 9]. However, because sensor elements and wiring tend to be concentrated near the contact surface, such sensors are structurally susceptible to mechanical damage and face difficulty in simultaneously achieving flexibility and large deformations. In addition, their performance is strongly affected by electromagnetic noise, humidity, and the material properties of the contact object, and thus complex compensation methods are often required to ensure stable operation [10].

Optical tactile sensors, by contrast, measure contact-induced deformations or changes in optical patterns using optical waveguides, optical fibers, or camera images, and offer advantages such as high spatial resolution, a wide dynamic range, and robustness against electromagnetic noise [8–10]. In particular, vision-based tactile sensors estimate contact by capturing the displacement of markers or variations in brightness patterns on the back side of a flexible skin with a camera, and then inferring contact geometry, force distribution, slip, and surface texture at high spatial density using image processing and machine learning techniques [11–14]. Nevertheless, in most cases, the processing pipeline is tightly coupled to a specific sensor design, so that designing a tactile sensor for a given application requires co-design of both the sensor hardware and the associated processing system. Furthermore, the sensing characteristics of a designed sensor cannot be verified without actual fabrication, making iterative design tuning toward the target application difficult. For these reasons, there is a strong demand for a unified system that can be applied in a scalable manner, from whole-body robot morphologies down to ultra-compact designs such as medical catheters.

1.1.2 Tactile sensors for medical applications

In recent years, there has been a growing demand in clinical practice for surgical techniques that minimize tissue incision in order to reduce the burden on patients. In response, a variety of medical devices, such as catheters and endoscopes, have been developed to facilitate minimally invasive surgery, and numerous surgical techniques for different pathologies using these devices have been proposed. In such minimally invasive surgery, not only the treatment of the target region but also the navigation of medical instruments along anatomical pathways requires highly precise and skillful manipulation. This is essential to minimize the physical burden on patients during insertion and positioning of the instruments.

To enable such advanced manipulation, it is crucial to provide the operator with rich information about the environment surrounding the medical device and thereby support intuitive operation. However, small surgical instruments such as catheters inherently do not afford intuitive manipulation. This is because current systems rely on indirect visualization methods, such as narrow field-of-view endoscopy or X-ray CT, and are unable to directly sense the target tissue in real time. Consequently, these procedures demand a high level of specialized knowledge and extensive clinical experience.

In light of these challenges, directly sensing and feeding back information about the interaction between medical devices and biological tissue has become an important requirement. In particular, medical tactile sensing technologies that can measure mechanical interactions—such as forces between tissue and instruments—are regarded as highly promising [15–17].

Recent studies have further proposed the use of tactile sensing for the diagnosis of malignant tumors by analyzing tactile information obtained from palpation of internal tissues [18, 19]. In these applications, compact tactile sensors suitable for use inside the human body have been developed, and in many cases, electrical sensing methods based on Micro Electro Mechanical Systems (MEMS) are employed. While such approaches enable compact sensor designs, they typically require expensive microfabrication processes and have limited capability to acquire tactile information that is spatially distributed over a wide area. As a result, the practical utility of these sensors remains restricted [16].

1.1.3 Ablation catheters

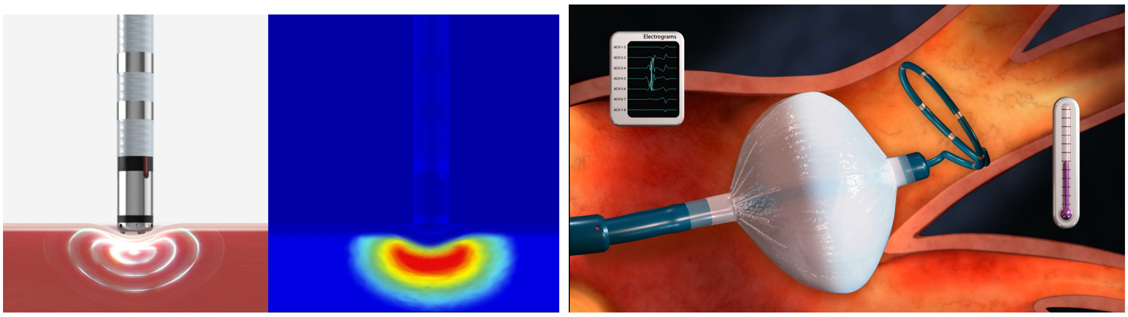
Atrial Fibrillation (AF) is the most common persistent arrhythmia and is known to be associated with an increased risk of stroke, heart failure, and all-cause mortality. Ablation catheters are therapeutic catheters used to treat arrhythmias such as AF by delivering thermal or non-thermal energy from the endocardial side to pathologic myocardial tissue that constitutes abnormal electrical conduction pathways. By thermally or non-thermally ablating this tissue, the myocardium is rendered necrotic and scar tissue is formed, thereby achieving electrical isolation. In particular, because ectopic activity originating from the pulmonary veins is a major trigger of AF, the current standard therapy is Pulmonary Vein Isolation (PVI), in which the ostia of the pulmonary veins are encircled by concentric scar lesions from the left atrial side to block electrical connections between the pulmonary veins and the main body of the left atrium [20, 21].

Ablation catheters can be broadly classified according to the mode of ablation. One category is point-by-point ablation, in which thermal or non-thermal energy is applied via point contact. In point-by-point ablation, the catheter tip is sequentially moved along the atrial endocardial surface, and ablation is performed at each contact point to form linear or circumferential lesion sets. Because the ablation outcome strongly depends on the stability of catheter-tip positioning and the contact conditions with the myocardium, it has been reported that even for the same operator, variability in lesion continuity and depth can easily occur, potentially affecting treatment outcomes [20–22].

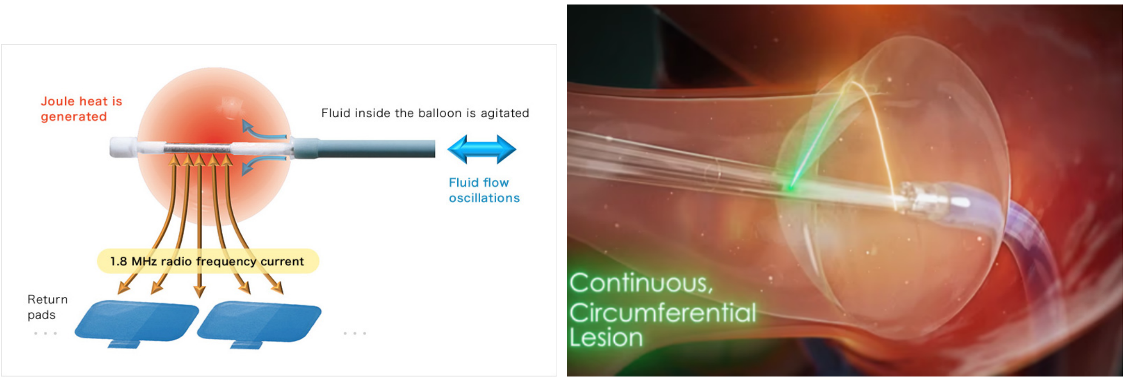
In contrast, balloon catheter ablation is a technique that enables circumferential ablation with a single catheter. In this approach, the balloon is inflated at the pulmonary vein ostium, its outer surface is brought into intimate contact with the myocardial tissue, and ablation is then performed to create a continuous scar. Balloon catheter systems include cryoballoon, which perform cryothermal ablation using liquid nitrogen or other refrigerants [23]; HotBalloon, which perform thermal ablation by heating the balloon’ s internal medium with radiofrequency current [24, 25]; and laser balloon, which perform ablation by irradiating from the inner side of the balloon with laser energy [26] [20–22].

Balloon catheter ablation can exploit the high deformability of the balloon to create extensive scar lesions and is considered capable of shortening procedure time while maintaining efficacy and safety comparable to standard point-by-point ablation. It is also reported to reduce dependence on the operator’ s level of experience and to facilitate relatively uniform isolation of the pulmonary veins. On the other hand, in cases with pulmonary vein anatomical anomalies or small-diameter veins, complete occlusion can be difficult to achieve, and issues such as incomplete ablation and pulmonary vein stenosis have been pointed out as remaining challenges [20–22, 27].

In balloon catheter-based cardiac ablation procedures, it has been shown that monitoring contact forces and mapping ablation sites strongly influences the overall procedural success rate [15–17]. Recent studies have further proposed the use of tactile sensing for the diagnosis of malignant tumors by analyzing tactile information obtained from palpation of internal tissues [18, 19]. In these applications, compact tactile sensors suitable for use inside the human body have been developed, and in many cases, electrical sensing methods based on MEMS are employed. While such approaches enable compact sensor designs, they typically require expensive microfabrication processes and have limited capability to acquire tactile information that is spatially distributed over a wide area. As a result, the practical utility of these sensors remains restricted [16].



(a) point-by-point ablation catheter [28] (b) cryoballoon catheter [29]



(c) HotBalloon catheter [30] (d) Laser balloon catheter [31]

Fig. 1.3: Modes of ablation catheters: (a) point-by-point ablation, in which ablation is performed via point contact at the catheter tip [28] (b) cryoballoon, in which the balloon is filled with liquid nitrogen and the contacted tissue is frozen to induce necrosis [29] (c) HotBalloon, in which the liquid inside the balloon is heated by radiofrequency (RF) current to ablate the tissue [30] (d) laser balloon, in which tissue ablation is performed by irradiating from inside the balloon with a laser [31].

1.2 Research purpose

Conventional medical tactile sensors have been structurally and fabrication-wise constrained to local point or line measurements, making it difficult to acquire tactile information over the entire surface of an instrument. Furthermore, in intrabody environments, restrictions on the field of view and the placement of force sensors make real-time, wide-area tactile sensing challenging. Against this background, the present work aims to realize a tactile sensing technology capable of accurately estimating wide-area contact geometries by newly designing and constructing a tactile balloon catheter based on marker-based tactile sensing. However, implementing tactile sensing in ultra-compact balloon catheters is technically demanding.

To address this, we propose a novel data-driven tactile sensing system, shown in Fig. 1.4, as a framework for constructing tactile sensors and simplifying their development. The proposed system is characterized by a tactile estimation model that is trained in a simulation environment to infer tactile data. Specifically, we first design the tactile sensor as a three-dimensional model and construct a virtual environment that integrates deformable-body simulation with optical simulation. Within this environment, deformations of the sensor skin and changes in optical markers induced by contact are automatically generated to build a large-scale supervised dataset. We then train a machine learning model for tactile estimation using this dataset and introduce a learning strategy to mitigate the sim-to-real (Sim2Real) gap that arises when deploying the model in real-world settings.

By implementing the above data-driven tactile sensing system, we aim to realize a tactile balloon catheter as a concrete embodiment of the proposed framework.

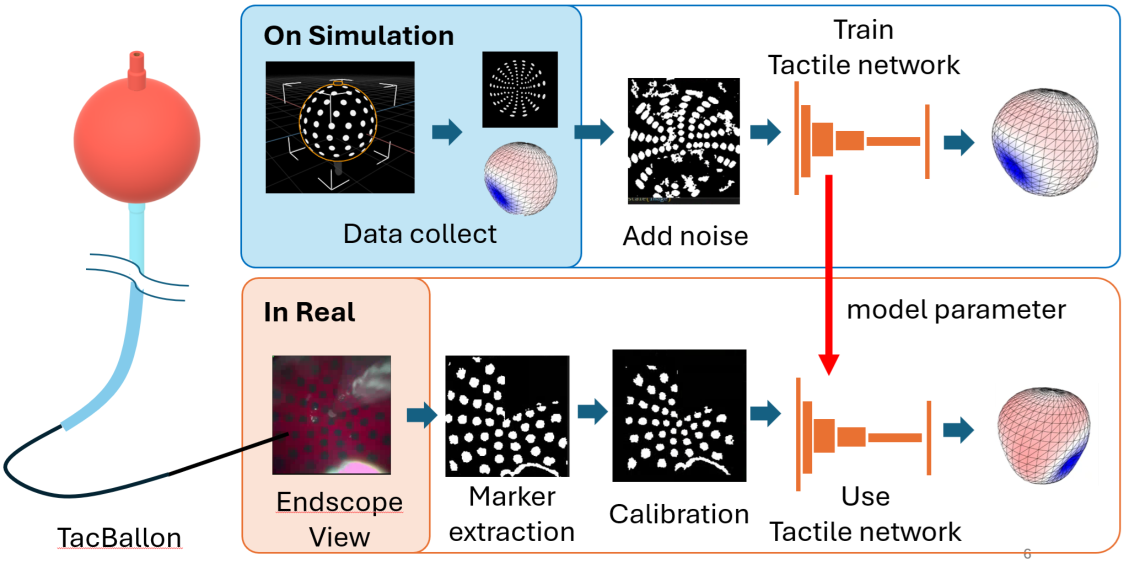


Fig. 1.4: System configuration: A machine-learning-based tactile estimation model is trained on a dataset collected through simulation, where noise is injected during training to mitigate discrepancies between the simulation and the real environment. Contact of the balloon catheter is then inferred by feeding the tactile estimation model with images in which optical markers have been extracted from endoscopic images acquired by an endoscope mounted inside the tactile balloon catheter.

1.3 Structure of the Research

This thesis proposes a data-driven, vision-based tactile sensing technique aimed at simplifying the construction of tactile sensing systems and advancing tactile sensing for minimally invasive surgery. As an application, we present a novel tactile balloon catheter and validate its effectiveness. Accordingly, the thesis is organized to develop the discussion step by step and systematically, covering background technologies, the proposed method, experiments, and discussion. An outline of each chapter is provided below.

Chapter1 (Introduction). We survey the landscape of tactile sensing and its medical applications. We clarify the advantages and technical challenges of vision-based tactile sensors adopted in the proposed method and summarize research trends in medical tactile sensing. We then identify the technical challenges involved in applying tactile sensing to balloon-catheter–type devices, position the problems addressed in this study, and articulate the significance of our approach.

Chapter2 (Related Works). We systematically review fundamentals of tactile sensing and prior work on medical applications. We outline the principles and characteristics of diverse tactile sensors—electrical, optical, fiber-optic, and fluidic—and describe the key features of vision-based tactile sensing. We also organize the challenges of applying tactile technologies to slender, compliant structures such as catheters and specify the technical problems targeted in this research.

Chapter3 (System Design, Fabrication, Simulation and Machine Learning). We describe the design philosophy and components of the tactile balloon catheter system developed in this work and detail the prototyping process. We then present the construction of the simulation environment used to generate training data for the tactile estimation model and provide a detailed description of the resulting machine learning dataset. The architecture of the learning model, as well as preprocessing and postprocessing procedures, are explained. The chapter concludes with the training process using simulated data, convergence characteristics, learning curves, and representative prediction examples.

Chapter4 (Evaluation). We evaluate the overall performance of the trained tactile estimation model and sensing system from three perspectives. First, we experimentally measure depth-estimation accuracy under point contact and discuss model performance for local balloon deformation. Second, using a rig with known shapes, we assess shape-estimation accuracy for area contact over the entire balloon and analyze discrepancies between the estimated mesh and ground-truth geometry. Third, we examine contact-area estimation for complex shapes, clarifying the effectiveness and limitations of the method with a view toward clinical use.

Chapter5 (Discussion). We synthesize the evaluation results and position the proposed data-driven tactile sensing system within the broader literature. In particular, we compare with prior vision-based tactile endoscopes, highlighting the characteristics, strengths, and limitations of our approach. We analyze the causes of accuracy degradation observed in different evaluations—such as uncertainty in out-of-FOV regions, differences between simulation and reality, and constraints in marker design—and outline technical issues for future improvement.

Chapter6 (Conclusion and Future Work). We summarize the contributions of this study and present directions for future research toward practical deployment and broader applications.

Chapter2

Related works

2.1 Overview

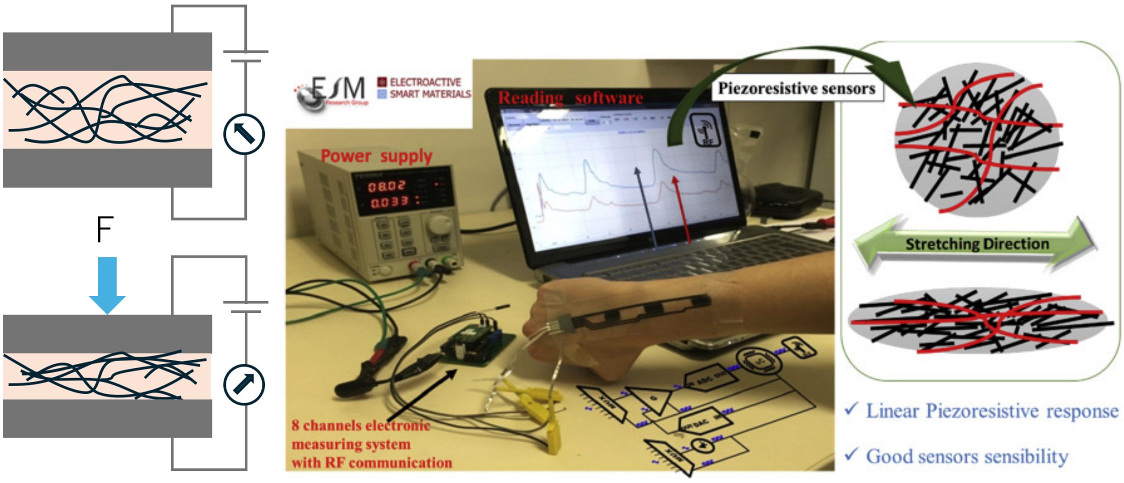
In this chapter, we review prior work on tactile sensing technologies, including those for medical applications, and clarify the context in which TacBalloon is situated. Tactile sensors can be broadly classified into approaches based on electrical, acoustic, fluidic, and optical methods. Each modality has its own strengths and weaknesses in terms of sensitivity, miniaturization, responsiveness, flexibility, and the ability to perform wide-area sensing, depending on its structure and sensing principle. Furthermore, as an application to medical devices, tactile sensors for catheters have been actively investigated. Owing to spatial constraints, tactile devices for medical use have thus far primarily adopted electrically based approaches.

2.2 Electrical approaches

In tactile sensors based on electrical approaches, external forces and contacts are converted into electrical signals to acquire tactile information. Representative methods exploit effects such as piezoresistance, resistance, capacitance, and piezoelectricity. In general, electrically based tactile sensors exhibit high sensitivity and good frequency response, while also consuming little power. On the other hand, because the sensing elements themselves are placed in close proximity to the contact surface, there is a durability issue in that damage to the elements or wiring can render sensing impossible. In addition, as these are electrical methods, sensing performance can be hindered by noise arising from the surrounding electromagnetic field or from the materials in contact. To ensure flexibility, it is necessary either to employ materials with special properties or to embed the sensor itself on the surface or inside an elastomer. Consequently, flexibility often trades off against structural durability and sensing accuracy. Furthermore, enhancing functionality and increasing density through MEMS entails substantial fabrication cost [8, 9].

2.2.1 Piezoresistance

Piezoresistive tactile sensors are composed of insulating polymers in which conductive fillers are dispersed, and detect mechanical deformation caused by external forces as



(a) (b) Goncalves *et al.* [32]

Fig. 2.1: Piezoresistive tactile sensor (a) Schematic of a piezoresistive tactile sensor: an external force compresses the polymer, bringing conductive fillers closer together and thereby reducing the resistance. (b) Piezoresistive tactile sensor composed of PVA and CNTs [32].

changes in electrical resistance. Typical conductive fillers include carbon black [33], carbon Nanotube (CNT) [32, 34, 35], and metal nanoparticles [36, 37]. Insulating polymers used as matrices include PolyDiMethylSiloxane (PDMS) [38], PolyUrethane (PU) [35], and PolyVinylAlcohol (PVA) [32]. Sensors that rely on the deformation of conductive fluids—such as liquid metals or ionic liquids—have also been proposed [39]. The change in resistance of such conductive composites is described by Eq. (2.1) [8, 9].

$$R = \frac{\rho L}{A} \tag{2.1}$$

R: resistance, ρ: resistivity,

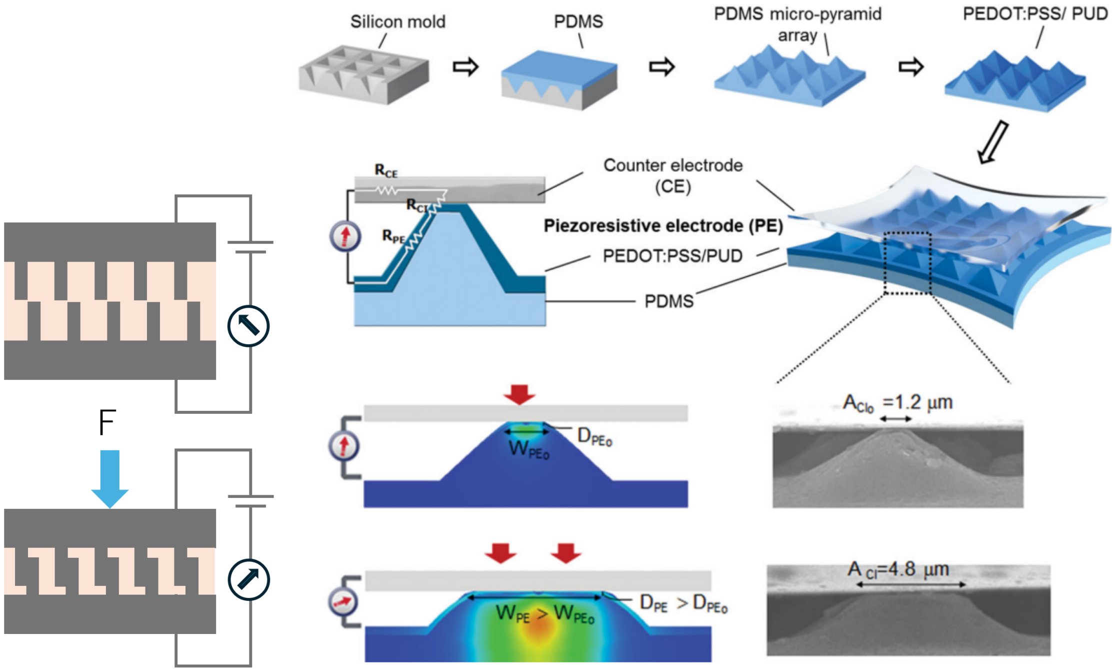
L: length of the resistor, A: cross-sectional area of the resistor

When an external force compresses the tactile sensor, L decreases, while A increases due to the Poisson effect, leading to a decrease in resistance. The resistivity ρ is determined by factors such as contact between conductive fillers dispersed in the insulating polymer and tunneling effects. Piezoresistive tactile sensors have the advantages of simple signal processing and robustness against electromagnetic noise. However, they tend to exhibit hysteresis, which degrades their frequency response characteristics [8, 9].

Gonçalves et al. developed an aqueous printable ink composed of PVA and CNT for fabricating piezoresistive sensors. PVA is a water-soluble polymer with stretchable mechanical properties, enabling the fabrication of piezoresistive sensors capable of withstanding strain levels of up to 3. Moreover, because the material is compatible with water-based printing, sensors of various shapes can be easily manufactured using this approach [32].

2.2.2 Resistive

Resistive tactile sensors utilize changes in contact resistance between conductors induced by external forces. Deformations in the microstructure at the conductor contact interface under external loading cause changes in resistance, which are used to detect the magnitude and direction of contact forces [40, 41]. The relationship between the resistance and the applied force is approximately proportional as expressed in Eq. (2.2).



(a) (b) Choong *et al.* [40]

Fig. 2.2: Resistive tactile sensor. (a) Schematic of a resistive tactile sensor. (b) Resistive tactile sensor: protrusions at the contact interface deform under external force, increasing the contact area and thereby decreasing the resistance [40].

$$R \propto F^{-\frac{1}{2}} \tag{2.2}$$

R: resistance, F: external force.

Resistive tactile sensors with such structures have the advantage of achieving high sensitivity even for small forces. However, due to issues such as reproducibility of the microstructure at the contact interface, they tend to exhibit hysteresis and drift [8–10].

Choong et al. developed a resistive tactile sensor by coating highly stretchable electrodes onto a compressible PDMS substrate patterned with microscale pyramidal structures. When an external force is applied, the pyramids deform, altering the contact area and geometry, which in turn changes the electrical resistance of the sensor [40].

2.2.3 Capacitive

Capacitive tactile sensors consist of two electrodes and a dielectric layer sandwiched between them, and measure external forces as changes in capacitance. By introducing microstructures into the electrodes and dielectric, sensor designs with enhanced sensitivity and flexibility have been proposed. In the case of a parallel-plate structure, the change in capacitance is expressed as in Eq. (2.3).

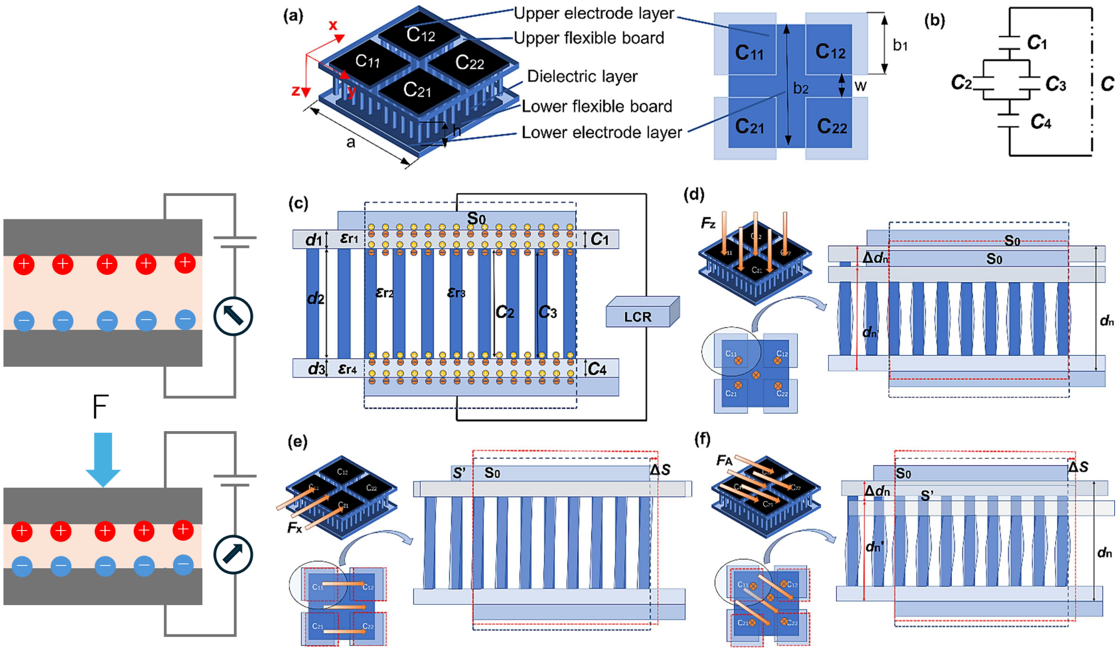
$$C = \frac{\epsilon_0 \epsilon_1 A}{d} \tag{2.3}$$

C: capacitance,

A: overlapping area of the plates, d: distance between the plates,

ϵ_0: permittivity of free space, ϵ_1: relative permittivity of the dielectric

Changes in d are used to measure normal forces through corresponding changes in capacitance, while changes in A can be exploited to measure shear forces. Capacitive



(a)

(b) Gong *et al.* [42]

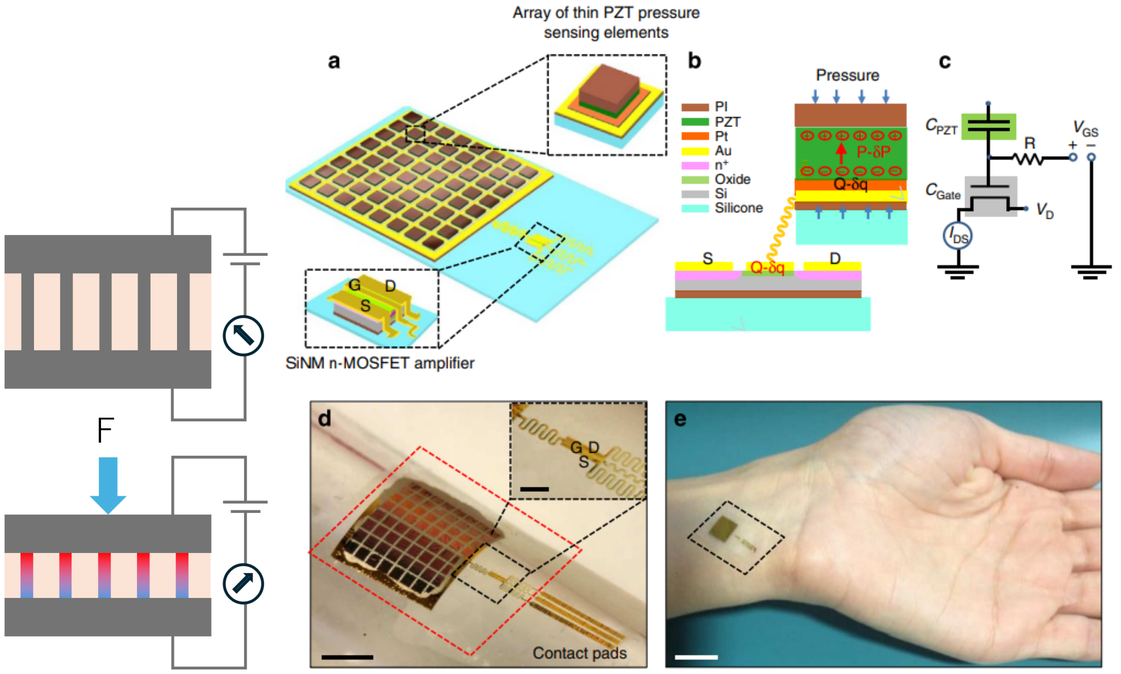
Fig. 2.3: Capacitive tactile sensor. (a)Schematic of a capacitive tactile sensor. (b)A capacitive tactile sensor composed of a microcylindrical array: the direction and magnitude of the applied force are calculated from differences in the output signals of a 2×2 array of capacitive sensor elements [42]

tactile sensors offer excellent frequency response and a wide dynamic range. However, they are susceptible to electromagnetic noise, and noise filtering requires relatively complex signal processing. They are also sensitive to humidity and may malfunction depending on the surrounding environment and the material of the contact object [8–10].

Gong et al. proposed a capacitive tactile sensor that employs a dielectric layer engineered with a micro-cylindrical structure. When an external force is applied, the air volume fraction between the cylindrical dielectric pillars decreases, resulting in a larger change in effective permittivity with respect to strain. This enhances the sensitivity and overall performance of the sensor. Furthermore, by analyzing the output signals from a 2×2 array of capacitive elements, the system can disentangle and measure both the magnitude and direction of the applied force [42].

2.2.4 Piezoelectric effect

Piezoelectric tactile sensors are based on the piezoelectric effect and measure tactile information using the piezoelectric potential generated when a piezoelectric material is mechan-



(a) (b) Canan *et al.* [43]

Fig. 2.4: Piezoelectric tactile sensor. (a) Schematic of a piezoelectric tactile sensor. (b) Thin-film piezoelectric tactile sensor: by leveraging the high sensitivity and fast response of piezoelectric tactile sensing, this device enables continuous monitoring of arterial blood pressure.

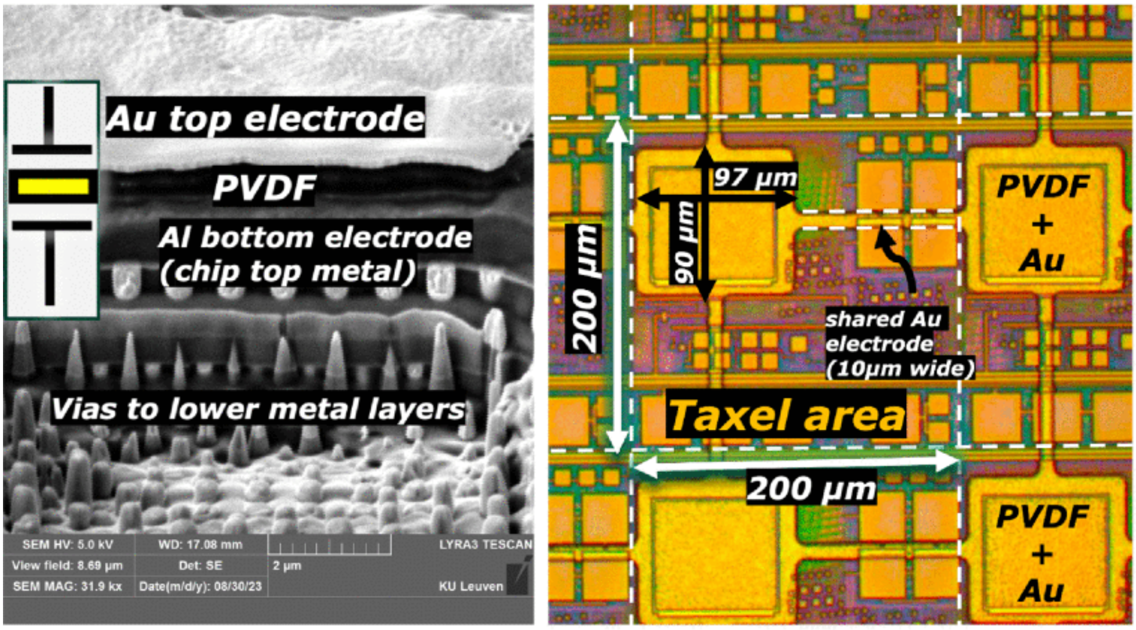
ically deformed. Sensitivity and mechanical flexibility depend on the material: inorganic piezoelectric materials such as ZnO [44] and GaN [45] have high Young's moduli, making them unsuitable for soft tactile sensors. For mechanically flexible piezoelectric materials, polymer-based materials such as PVDF are often used [46]. Because piezoelectric tactile sensors generate a piezoelectric potential only at the instant of mechanical deformation, they exhibit excellent high-frequency response and are advantageous for vibration measurement. However, owing to their high internal resistance, they cannot measure static contacts. Moreover, piezoelectric materials also exhibit pyroelectricity, and sensor temperature can therefore influence the measurement results [8, 9].

Canan et al. proposed a piezoelectric tactile sensor based on the piezoelectric effect of lead zirconate titanate (PZT). By leveraging the high piezoelectric coefficient of PZT and the fast dynamic response inherent to piezoelectric tactile sensors, they developed a device capable of continuously measuring subtle deformations induced by slight variations in arterial pressure [43].

2.2.5 MEMS

Among the advancements in the aforementioned electrical tactile sensing methods, MEMS technologies have become increasingly important for miniaturizing tactile sensors. MEMS technology uses semiconductor fabrication processes such as photolithography, etching, and thin-film deposition to create micro-scale mechanical structures, enabling the production of high-precision sensing elements on the millimeter to sub-millimeter scale. This technology has facilitated proposals for high-density, high-resolution, and multifunctional tactile sensors that exploit electrical effects. However, MEMS microstructures are generally fragile and prone to damage in environments involving large deformations. In addition to fabrication costs, design costs associated with miniaturization—such as wiring layouts and structural design—are also non-negligible [8, 9].

Alea *et al.* proposed a tactile sensor implemented on a chip by integrating a PVDF-based piezoelectric sensor array with a front-end signal processing chip. The sensor and chip were implemented using CMOS technology, achieving dimensions of $200\,\mu m \times 200\,\mu m$ per cell [47].



(a) (b)

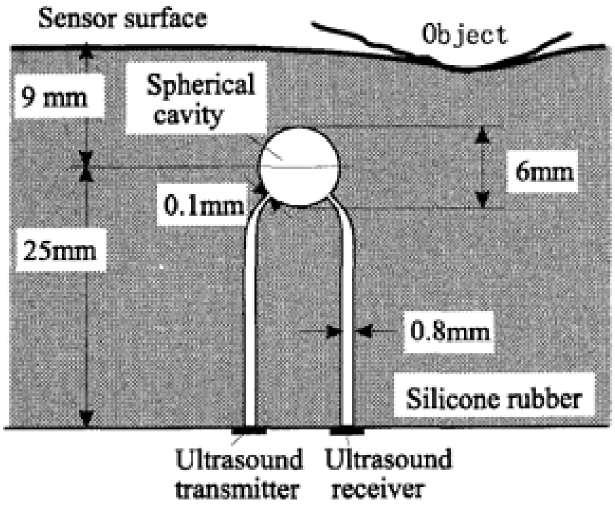
Fig. 2.5: Tactile Sensor Using MEMS (a) Cross-sectional SEM image of the tactile sensor (b) Chip micrograph of the tactile sensor implemented on a CMOS chip [47]

2.3 Acoustic-based tactile sensors

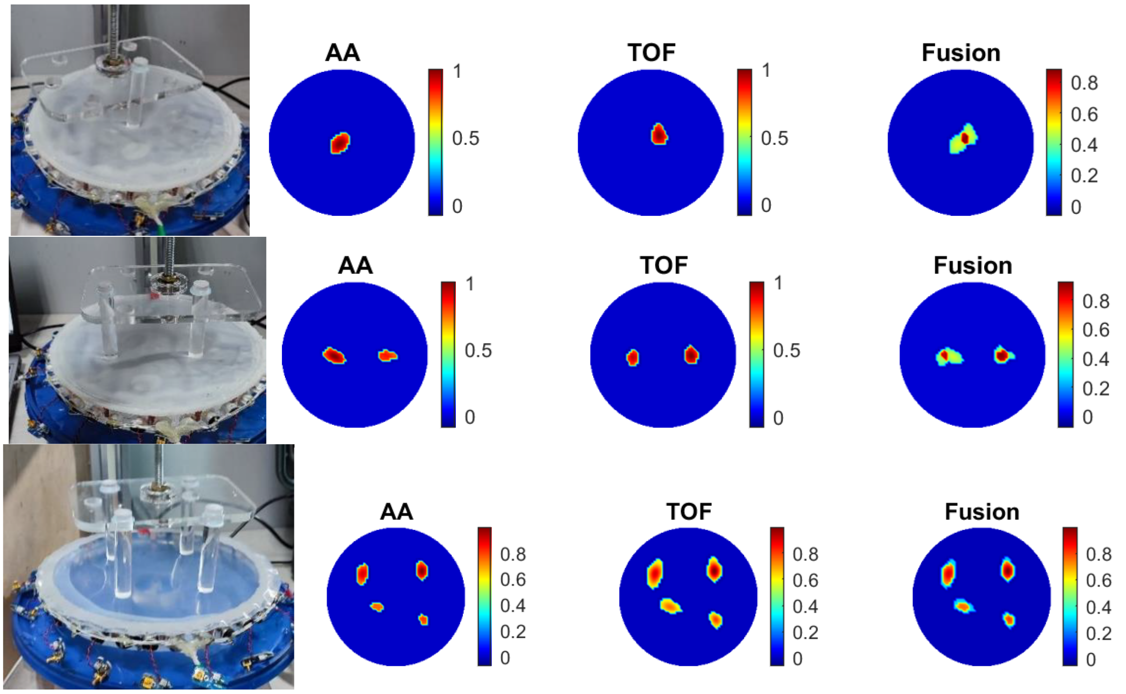
Acoustic-based tactile sensors estimate contact, pressure, and the mechanical properties of contacted objects by measuring changes in the characteristics of acoustic waves propagating inside or on the surface of an elastic body. Typically, ultrasonic or audible elastic waves are excited within the sensor, and tactile information is obtained by observing changes in resonance frequency, attenuation, and phase or amplitude spectra induced by contact.

A key advantage of acoustic tactile sensors is that they exploit the wave field generated in the fluid inside the sensor body and therefore do not necessarily require complex electrodes or wiring at the contact surface. Moreover, by tuning the wavelength and frequency band of the acoustic waves, it is possible to adjust sensing characteristics, such as sensitivity not only to surface contact but also to information in the depth direction to some extent. On the other hand, resonance characteristics strongly depend on the properties of the medium and the contact conditions, making compensation for environmental fluctuations, temperature variations, and aging effects crucial. In addition, in flexible structures with complex propagation paths, it has been pointed out that identifying the relevant modes is challenging.

The acoustic tactile element proposed by Shinoda *et al.* consists of a hollow elastic body and ultrasonic transducers, as shown in Fig. 2.6a. When the internal cavity of the elastic body is deformed, its acoustic resonance characteristics change, and the direction and magnitude of the external force are estimated from these changes. This method exploits the fact that the eigenmodes of the spherical cavity are slightly perturbed by the external force, and that this perturbation appears in the frequency response [48]. Soleimani *et al.* proposed an ultrasonic-tomography-based tactile skin that reconstructs the spatial distribution of acoustic attenuation using multiple ultrasonic transducers arranged around the periphery, thereby enabling wide-area contact detection and pressure estimation. In this approach, the entire sensor surface is treated as a single continuous medium, and the contact location and distribution are reconstructed from changes in the propagation characteristics of the acoustic field within the medium [49].



(a) Shinoda *et al.* [48]



(b) Soleimani *et al.* [49]

Fig. 2.6: Acoustic tactile sensors. (a) Tactile sensor based on changes in frequency response: contact is estimated from changes in acoustic characteristics induced by deformation of a spherical cavity [48]. (b) Tactile sensor based on ultrasonic tomography: the distribution of tactile stimuli is obtained by reconstructing an attenuation map of acoustic waves from changes in their propagation characteristics [49].

2.4 Fluidic tactile sensors

Fluidic tactile sensors estimate externally applied forces and deformations by measuring changes in the pressure or flow rate of a fluid enclosed within the sensor body. Typical working fluids include air, water, saline solution, and liquid metals, which are filled into hollow chambers embedded inside compliant structures [10].

Goshtasbi *et al.* modeled fluid–structure interactions and proposed a method that estimates contact location and load by training a machine learning model on internal pressure measurements [50]. Zou *et al.* proposed a sensing strategy that can be retrofitted to existing soft fluidic actuators. In their approach, the deformation state of the actuator and the shape, stiffness, and surface roughness of the contacted object can be inferred solely from measurements of the drive-side fluid pressure and flow rate. By modeling the input–output relationship of the actuator, tactile information can thus be estimated without embedding additional sensing elements into the structure, enabling tactile functionality to be easily added to existing fluid-driven devices [51].

Instead of measuring pressure directly, alternative designs treat the fluid within the flow channel as a conductive medium and read out pressure-induced cross-sectional changes as variations in electrical resistance, effectively realizing a piezoresistive sensor. Several tactile sensors employing conductive fluids—such as low-melting-point liquid metals or ionic liquids that remain liquid at room temperature—have been proposed. Yao *et al.* presented a thin, wearable tactile sensor in which a gallium-based liquid metal is encapsulated in cavities inside an elastomer. Changes in the cross-sectional area of the internal cavity due to bending or contact are detected as variations in conductive resistance and used as tactile signals [52]. Wang *et al.* proposed a tactile sensor, illustrated in Fig. 2.7, that employs three liquid-metal-filled microchannels formed by 3D printing; their device simultaneously detects tactile stimuli and temperature changes by reading out resistance changes in the three microchannels [39]. In contrast, Kim *et al.* proposed a soft pressure sensor that uses physiological saline as a conductive fluid. Owing to its high biocompatibility, saline offers advantages over liquid metals in terms of safety and thermal stability [53].

Su *et al.* proposed the method shown in Fig. 2.8, in which tactile information is estimated from changes in both the pressure of a weakly conductive fluid and the impedance between electrodes induced by fluid deformation. When the elastomer deforms upon contact with an object, the internal conductive fluid is deformed, leading to changes

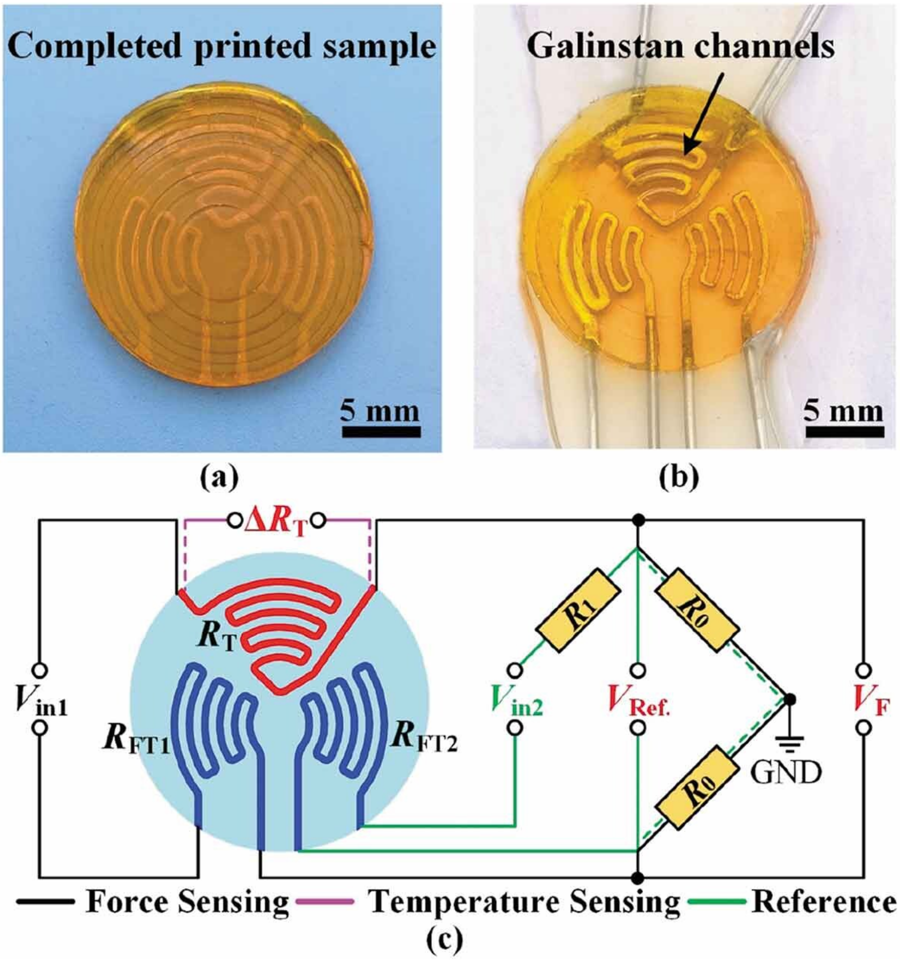


Fig. 2.7: Tactile sensor based on conductive fluid. (a) Configuration of the microchannels. (b) Encapsulated liquid metal: deformation of the microchannels leads to deformation of the liquid metal. (c) Sensing circuit: variations in resistance caused by deformation of the liquid metal in the three microchannels and by changes in the liquid metal temperature are read out for sensing [39].

in pressure. At the same time, the deformation of the conductive fluid alters the conductive paths between electrodes placed on the chamber surface, and the resulting changes in impedance are used to measure tactile stimuli [54].

A major advantage of fluidic tactile sensors is that they can maintain high flexibility of the overall structure while allowing the electronic circuitry and pressure sensors to be located remotely from the contact region. This is beneficial in environments where placing metals or electronic components at the contact interface is undesirable, such as in strong electromagnetic environments or medical applications. Furthermore, by designing the internal fluid domain and hollow geometry, local contacts can be made to actively induce specific pressure distributions, from which the contact location and contact geometry can

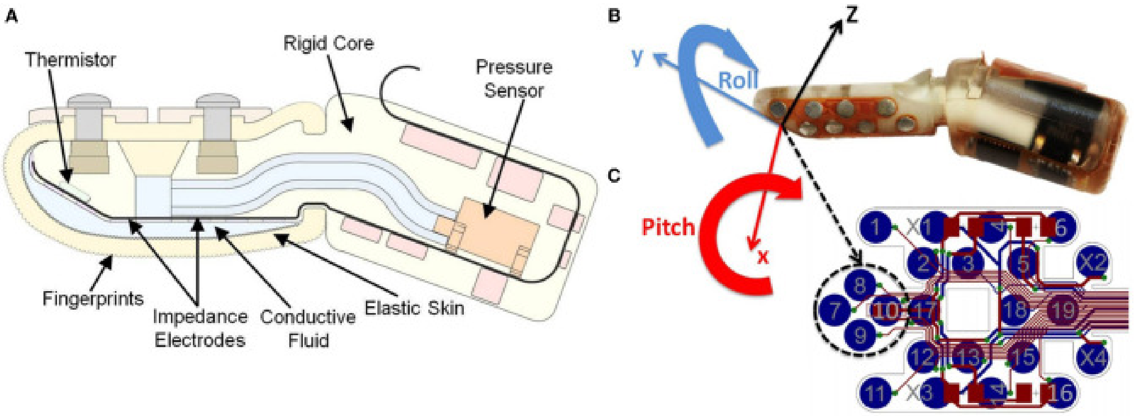


Fig. 2.8: Tactile sensor based on pressure and impedance of a conductive fluid: tactile stimuli are measured from impedance changes arising from variations in the conductive paths between electrodes as the conductive fluid deforms [54].

be inferred.

On the other hand, because fluids possess viscosity and inertia, fluidic sensors exhibit dynamics that differ from those of purely electrical sensors in terms of response speed and hysteresis. In particular, in long tubes or complex flow paths, propagation delays and losses of pressure waves become significant, making modeling and compensation indispensable in applications that require fast tactile responses. Designs that require auxiliary components such as pumps and valves also face challenges in miniaturizing and integrating the overall system. When liquid metals are used, additional issues must be considered, including changes in material properties due to oxidation and temperature variations, as well as embrittlement caused by corrosion of other metals.

2.5 Optical tactile sensors

Optical tactile sensors are sensing devices that convert mechanical contact and deformation into tactile information by reading changes in optical propagation or image patterns. Optical tactile sensors offer high spatial resolution and a wide frequency response range. They also have advantages such as high durability, flexibility, and robustness against electromagnetic and thermal noise. On the other hand, they typically require relatively high power consumption and a computing system capable of performing complex processing.

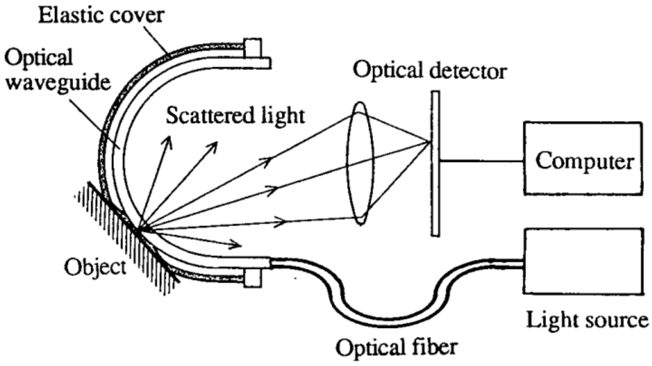
A basic optical tactile sensor consists of a light source, an optical medium (e.g. , waveguides, elastomers, or optical fibers), and a photodetector (e.g. , photodiodes or cameras). When the sensor surface deforms upon contact, the internal optical path length, conditions for transmission, reflection, and scattering, or visual patterns change. By detecting and analyzing these changes, it is possible to estimate contact location, contact force, surface geometry, and other tactile features. Optical tactile sensing approaches can be broadly classified into methods based on optical waveguides and optical fibers, and vision-based methods that acquire tactile information as images using cameras [8–10].

2.5.1 Optical waveguide tactile sensors

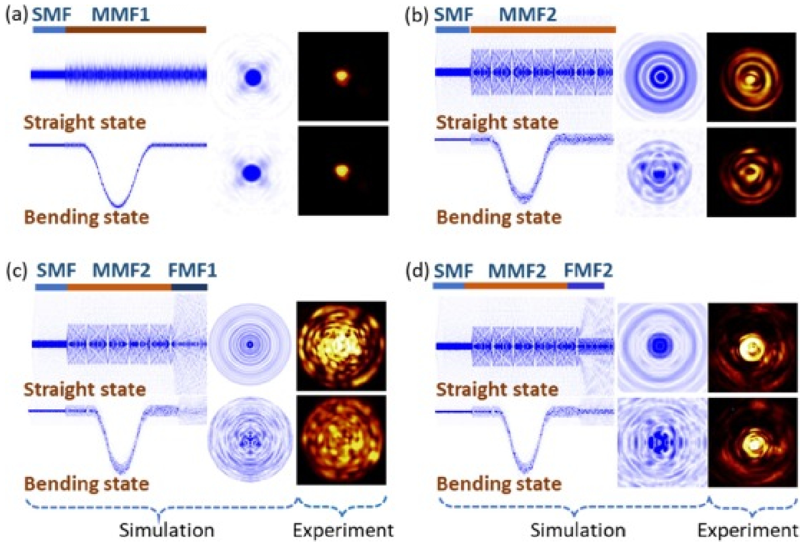
Optical waveguide tactile sensors exploit the properties of optical waveguides such as optical fibers, and measure contact-induced changes in light intensity, wavelength, transmission, and propagation path arising from deformation or strain of the waveguide [55–58].

Maekawa *et al.* proposed a tactile sensor in which a hemispherical optical waveguide is formed inside a fingertip-shaped structure. When a deformed elastomer is pressed against the surface of the optical waveguide, the reflection conditions at the waveguide surface change. As a result, part of the light that propagates inside the waveguide via repeated total internal reflection is scattered at the contact region, and tactile information is obtained by detecting this scattered light with photodetectors [59].

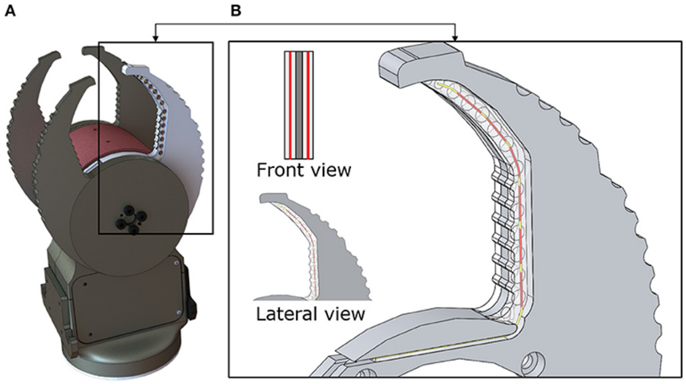
Luca *et al.* proposed a robot hand that uses optical fibers integrated with fiber Bragg grating (FBG) strain sensors. A fiber Bragg grating is a device in which a periodic modulation of the refractive index is formed in the core of an optical fiber. This periodic refractive index modulation acts as a diffraction grating and reflects only a specific wavelength component of the incident light. When a broadband spectrum is launched into a fiber Bragg grating, light at a specific wavelength (the Bragg wavelength) is reflected. The



(a) Maekawa *et al.* [59]



(b) Liu *et al.* [58]



(c) Luca *et al.* [57]

Fig. 2.9: Optical-waveguide tactile sensors. (a) Optical waveguide contact sensor: when the deformable skin is pressed and comes into contact with the optical waveguide, the reflection coefficient at the waveguide surface changes, total internal reflection is locally broken, and the leaked light is used to detect contact . (b) Intensity modulation of light induced by bending of an optical fiber . (c) Gripper-type tactile sensor using fiber bending: the red lines indicate optical fibers forming the tactile sensor, and tactile information is measured from the modulation of light caused by fiber bending .

Bragg wavelength shifts as a function of the temperature and strain applied to the optical fiber. In the work by Luca *et al.* , the sensor output was derived solely from strain under constant temperature conditions, and they demonstrated that both the size and stiffness of the grasped object can be accurately estimated using this approach [57].

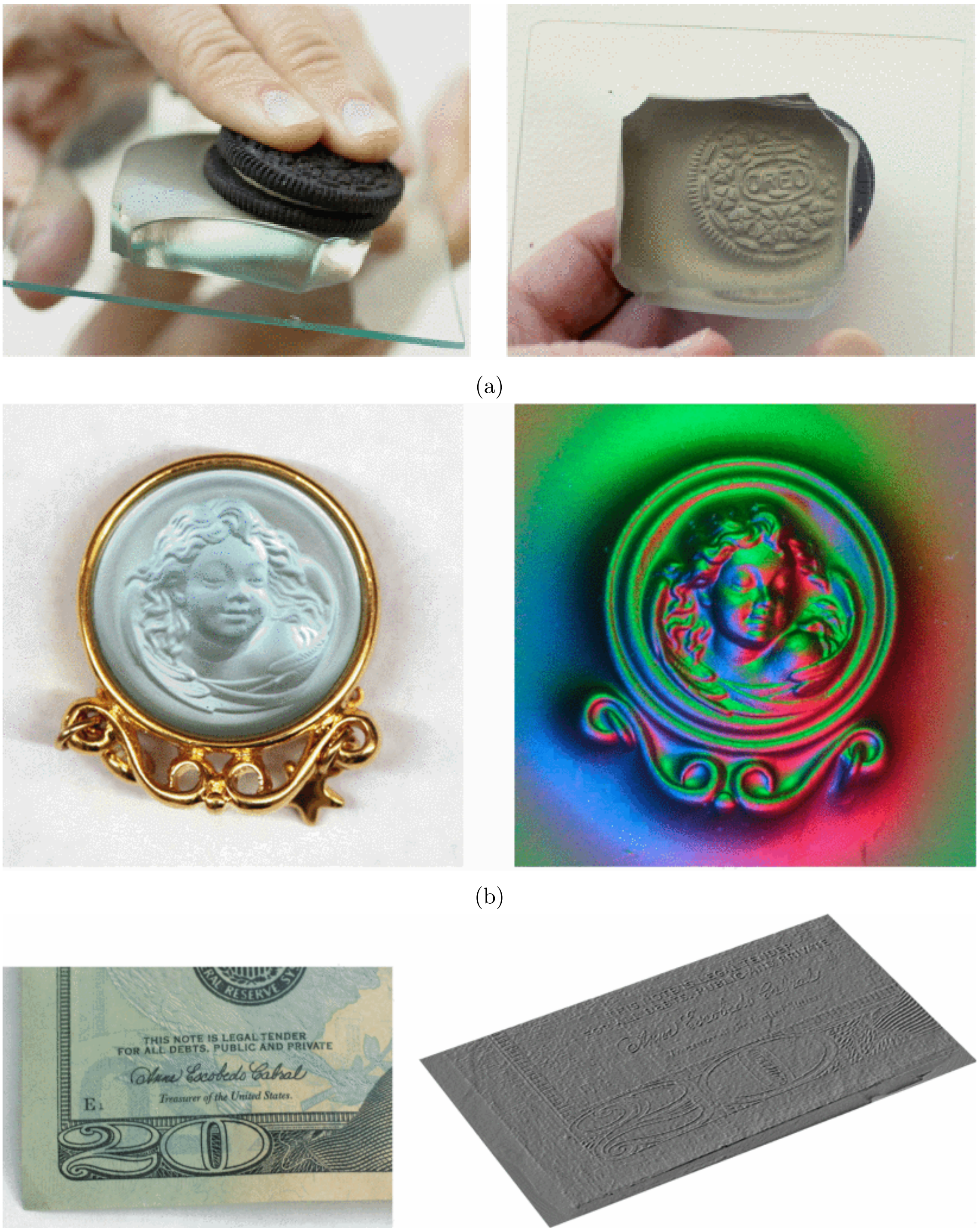
2.5.2 Vision-based tactile sensors

Vision-based optical tactile sensors use cameras to acquire tactile information as images. They combine a compliant skin with light sources and a camera, and capture images of marker displacements or brightness pattern changes on the back surface of the skin induced by contact. A key feature of this approach is that, by applying image processing and machine learning, it can estimate a wide range of tactile information, including contact geometry, force distribution, slip, and surface texture.

Vision-based tactile sensing approaches can be broadly divided into two categories: the GelSight method, which estimates contact from changes in reflected intensity of RGB light components, and marker-based methods, which track the motion of markers printed on or embedded in a flexible skin.

The GelSight sensor proposed by Johnson *et al.* consists of a transparent acrylic plate, a flexible skin covering its surface with a reflective coating, and internal illumination and a camera. As illustrated in Fig. 2.10b, deformations on the back surface of the skin caused by contact are captured as images of light intensity distributions. High-density tactile information is then extracted from these intensity distributions [11]. In particular, as shown in Fig. 2.10c, GelSight is well suited for visualizing fine surface textures formed by microscopic asperities with high resolution. However, the reflective coating on the skin surface is prone to wear, which limits durability. In addition, the sensor surface cannot undergo large deformations due to its structure. Furthermore, the arrangement of the camera and light sources is restricted, making it difficult to freely design arbitrary sensor geometries [60].

Marker-based tactile sensing methods consist of a flexible skin embedded with optical markers and a camera that tracks these markers to estimate deformation. The TacTip family adopts the design shown in Fig. 2.11. Numerous pin-shaped optical markers are arranged on a compliant skin, and an internal camera is placed so that it can observe the markers. Marker positions are extracted from the camera images by image processing



(a)

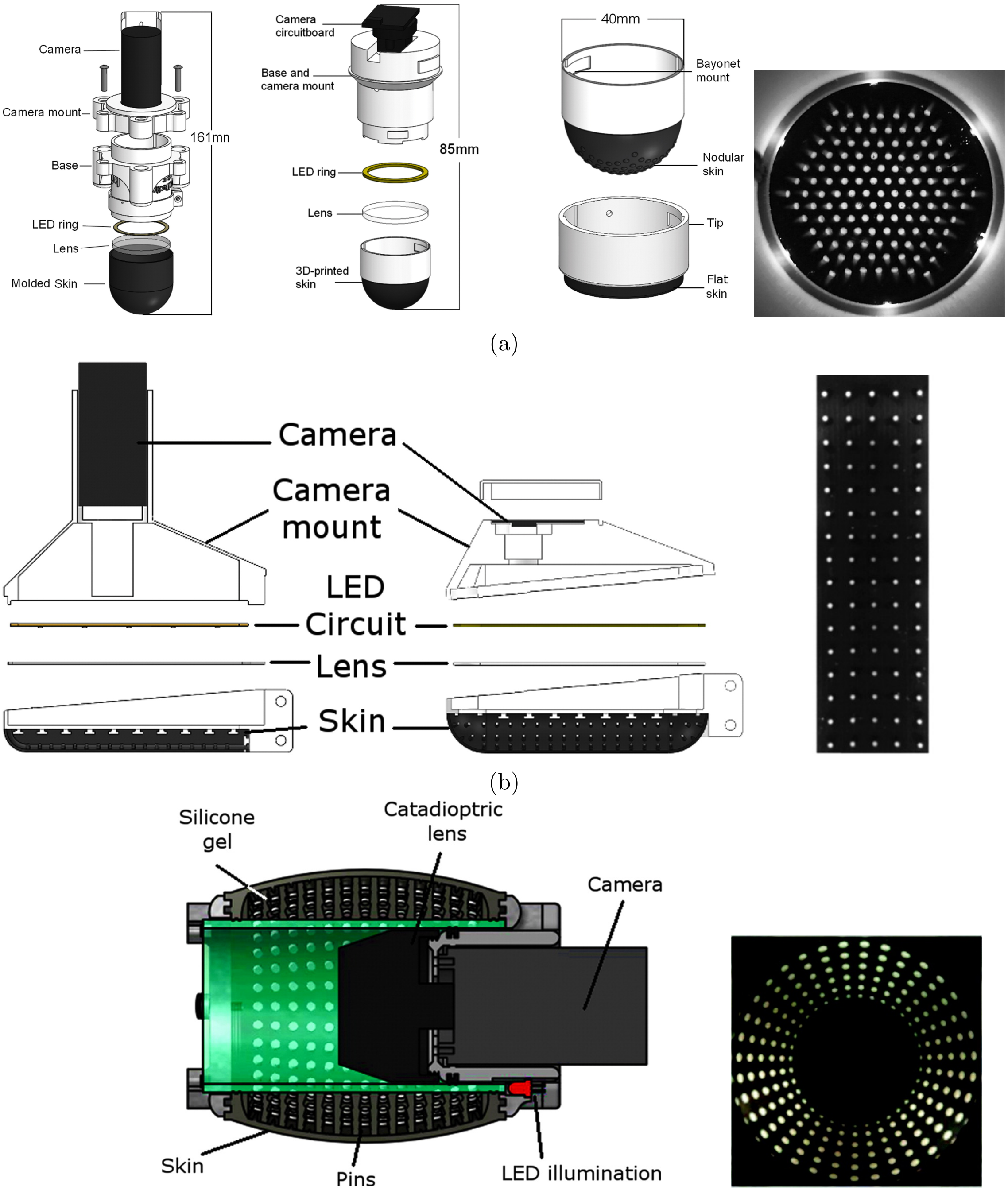
(b)

(c)

Fig. 2.10: GelSight tactile sensor. (a) Reflective-membrane elastomer on an acrylic plate, a cookie pressed against it, and the resulting deformation. (b) Deformation of the reflective membrane illuminated by blue light from the lower left and red light from the lower right. (c) Surface texture of a 20-dollar bill captured at high resolution with GelSight [11].

and converted into three-dimensional coordinates, enabling high-resolution measurement of skin deformation. As illustrated in Fig. 2.11, marker-based tactile sensors allow relatively high flexibility in marker layout and internal geometry design, and such design choices can be used to tune sensitivity to local curvature, shear deformation, and slip direction. Their effectiveness has been demonstrated in various robotic tasks, such as detecting slip and shear forces [12, 13, 61].

Moreover, marker-based tactile sensors are highly scalable. Duong *et al.* developed a large-area soft-skin tactile sensor that covers an entire robot, as shown in Fig. 2.12. They proposed a method that employs an internal stereo-camera-based vision system to measure wide-area contact distributions in a single shot [14].



(a)

(b)

(c)

Fig. 2.11: TacTip family: deformation of the black skin is inferred from the motion of white optical markers. Left: designs of marker-based tactile sensors. Right: camera views. (a) TacTip. (b) TacTip-M2. (c) TacCylinder [12].

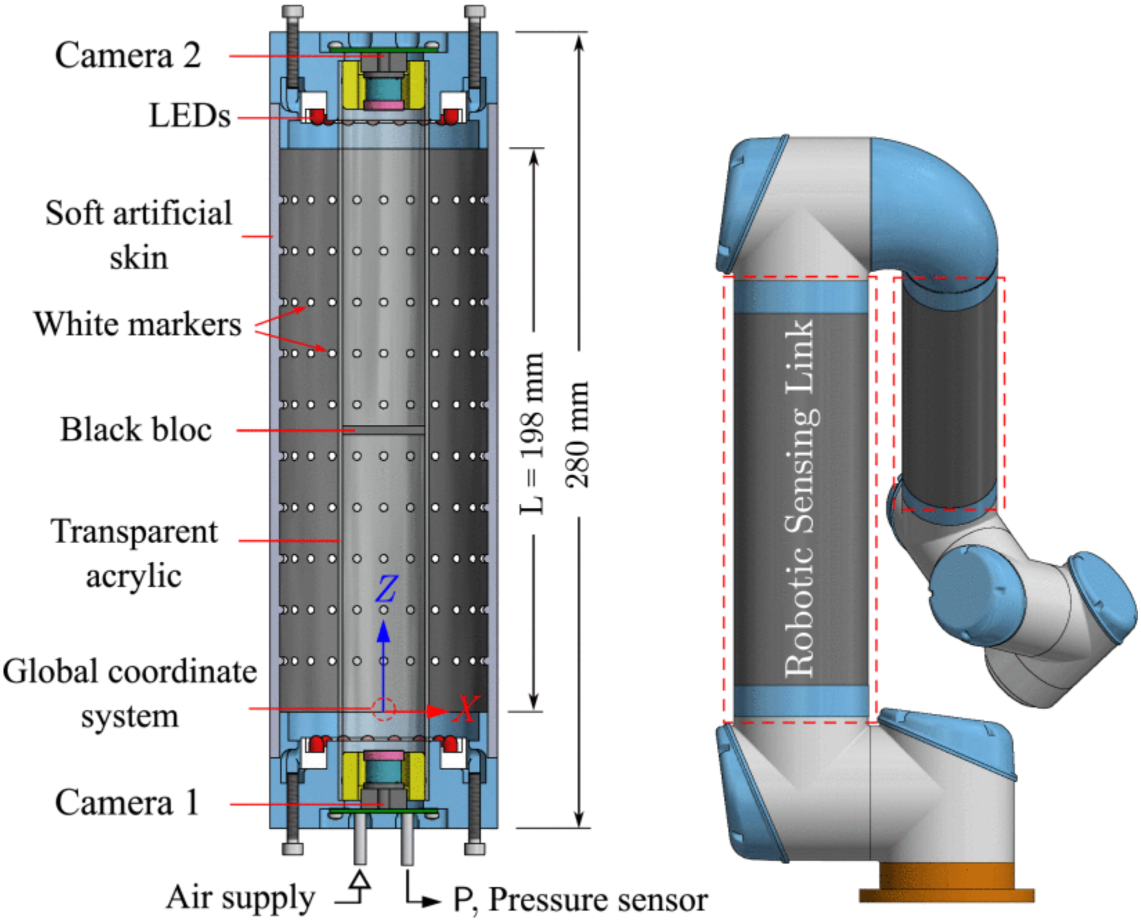


Fig. 2.12: Large-area tactile sensor covering an entire robot: tactile information is estimated from the three-dimensional coordinates of white optical markers measured by stereo cameras [14].

2.6 Medical tactile sensors

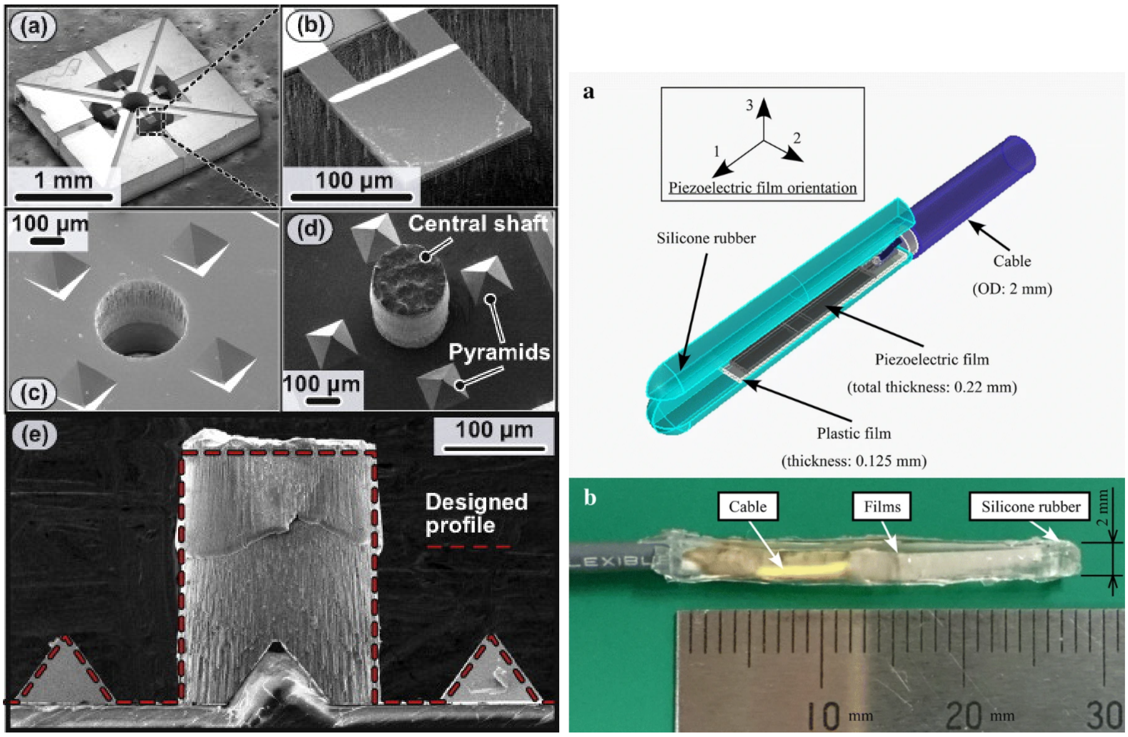
2.6.1 Catheter tactile sensors

The importance of tactile sensors in the medical domain has been rapidly increasing in step with advances in minimally invasive surgery. In particular, in endoscopic procedures and the manipulation of intravascular devices, surgeons lose the direct haptic feedback that was available in conventional open thoracic or abdominal surgery. As a result, tactile sensing technologies that can measure interactions with tissue have become indispensable. Tactile sensors used in clinical settings must prioritize safety with respect to biological tissues, requiring high flexibility, while simultaneously being constrained to miniature geometries that can be inserted into blood vessels or internal organs. For example, in intravascular catheters widely used in cardiac interventions, the typical outer diameter at the tip is on the order of 3. 5–4 mm, and even in the largest devices it does not exceed approximately 8–10 mm. Accordingly, tactile sensor design is subject to extremely stringent dimensional constraints, with an upper limit of about 8 mm, and in actual clinical use even thinner profiles are desired [17, 62].

Most medical tactile sensors proposed to date employ electrically based sensing principles, including piezoelectric, capacitive, and resistive mechanisms. These approaches are characterized by high sensitivity and relative ease of miniaturization via MEMS technologies. In addition, a variety of methods based on optical fiber sensing and other principles have been proposed. However, for clinical application, such sensors must simultaneously satisfy multiple requirements, including miniaturization, flexibility, durability, and resistance to electromagnetic interference [17, 62].

Clinical studies have demonstrated that real-time acquisition of tactile information can lead to improved treatment outcomes. A representative example is ablation catheter procedures in which contact-force sensors are employed. In these studies, ablation catheters equipped with a force sensor at the tip measure and control the contact force between the catheter and the myocardium during the procedure. This has been reported to improve the quality of lesion formation and reduce the recurrence of atrial fibrillation after the procedure [15–17].

Tactile sensors based on the piezoelectric effect constitute one of the major approaches in medical applications [63]. With advances in MEMS technology, it has become possible

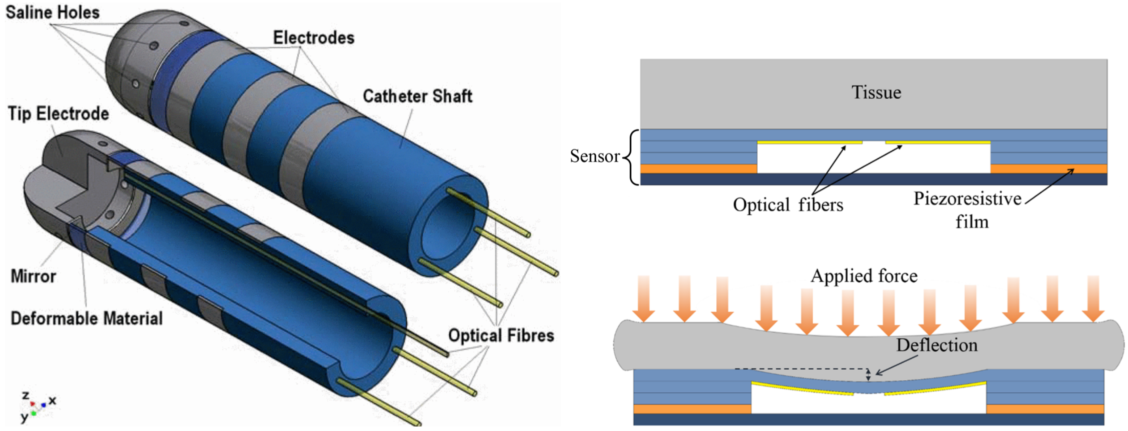


(a) Nguyen *et al.* [65] (b) Takashima *et al.* [70]

Fig. 2.13: Electrical tactile catheters. (a) Cantilever-type piezoelectric contact sensor fabricated using MEMS technology [65]. (b) Tactile sensor constructed by sandwiching a piezoelectric film between silicone rubber layers [70].

to fabricate micro-scale cantilever structures, such as those shown in Fig. 2.13a. Various methods have been proposed that achieve miniaturization, thin form factor, high-speed response, linearity, high sensitivity, and multi-axis detection of very small forces ($<$ 1 N) [64–68]. Other approaches have also been proposed in which piezoelectric films are sandwiched between flexible materials, allowing load measurement while conforming to compliant structures [69, 70]. However, in methods where electrically based sensor arrays share the same surface as the contact interface, or where the sensors are embedded directly at the contact surface, there is an inherent trade-off: structural flexibility is reduced, and the sensor elements and wiring become mechanically vulnerable, so that damage can easily render the device inoperable. This is particularly problematic in catheter structures, which are subjected to strong bending and torsional loads, making it difficult to ensure long-term reliability.

Furthermore, when such electrically based sensor arrays share the same surface as the contact interface, or when the sensor arrays are directly embedded at the contact surface, not only is flexibility reduced, but the concentration of sensors and wiring at the inter-



(a) Polygerinos *et al.* [72]

(b) Bandari *et al.* [75]

Fig. 2.14: Fiber-optic tactile catheters. (a) Tactile catheter based solely on optical fibers [72]. (b) Hybrid tactile sensor combining optical fiber bending with a piezoelectric element [75].

face increases the likelihood of mechanical damage. The use of MEMS-based fabrication processes also drives up manufacturing costs. In addition, electrically based measurement techniques are susceptible to electromagnetic influences, particularly in radio-frequency (RF) environments such as those used for ablation.

Tactile sensing methods based on optical fibers offer advantages such as small size, light weight, and high immunity to electromagnetic interference, making them suitable for use in combination with RF devices. In fiber-optic tactile sensors, configurations such as those shown in Fig. 2.14a have been proposed, in which contact forces are estimated by analyzing changes in light intensity or interference patterns induced by deformation or displacement of a diaphragm at the tip [71–73]. Hybrid tactile sensors that combine piezoelectric elements with optical fibers, as illustrated in Fig. 2.14b, have also been reported [74, 75].

Despite their advantages in terms of high sensitivity, miniaturization, and fast response, both electrical and fiber-optic tactile sensors face structural limitations when it comes to spatially resolving extensive contact with biological tissue. In practice, they are generally restricted to point or line contact measurements. Electrical sensors often lack sufficient flexibility due to embedded components, or have limited deformation capability. Fiber-optic sensors are likewise constrained in their allowable deformation because they are highly sensitive to changes in reflection and intensity. Consequently, although these devices are capable of measuring very small forces, their measurement range is limited, and they cannot accommodate larger forces.

2.6.2 Colonoscope tactile sensors

In colonoscopy, it has been clinically reported that flat polyps and non-polypoid tumors may be overlooked even when they lie within the endoscopic field of view, due to poor visual contrast. On the other hand, pathological studies have shown that tumor tissue is generally harder and exhibits mechanical properties distinct from those of normal mucosa, and tissue stiffness has been reported as a potential diagnostic marker for lesions [76]. Against this background, approaches have been proposed that do not rely solely on optical endoscopic images, but instead augment the endoscope with tactile sensing capabilities to quantitatively measure mechanical interactions with the intestinal wall and thereby detect or highlight lesion sites.

Kim *et al.* proposed a colonoscope tactile sensor that employs a donut-shaped balloon, as shown in Fig. 2.15. A reflective membrane is provided over a portion of the donut-shaped balloon surface, and an optical unit is inserted into the central opening of the balloon. RGB LEDs illuminate the reflective membrane independently, and the membrane is imaged by an endoscopic camera. The measurement principle is similar to that of the GelSight method: deformation of the balloon—and hence deformation of the reflective membrane—is projected onto RGB intensity maps and acquired as texture information. The device is designed with a minimum diameter of 25 mm and can accommodate diameters up to 90 mm by inflating the balloon [77].

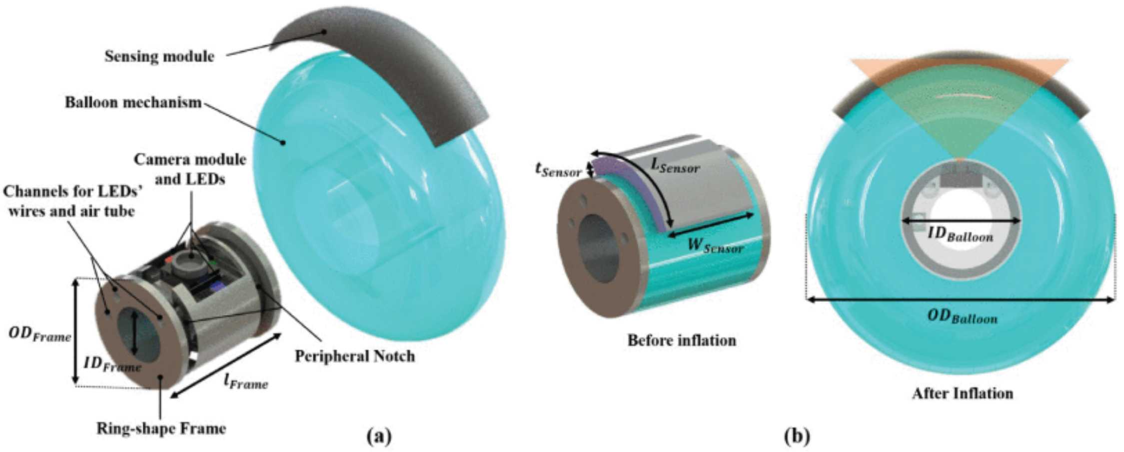


Fig. 2.15: GelSight-type colonoscopy tactile sensor. (a) Configuration consisting of an optical unit and a donut-shaped balloon with a reflective membrane. (b) Balloon shapes before and after inflation, and the camera field of view after inflation [77].

Chapter3

Data-driven tactile balloon catheter

3.1 Overview

Marker-based vision tactile sensors estimate tactile information by reconstructing the three-dimensional shape of a deformable skin from changes in marker positions using image processing. However, the image-processing pipeline depends on design factors such as marker layout and camera configuration, and thus methods based on machine learning have been proposed. In these approaches, the relationship between deformation and the resulting image changes is learned only after a physical sensor has been fabricated, which increases the cost of sensor development. Therefore, it is desirable to generate training data by simulation and to carry out model training in a simulation-based manner.

In this chapter, we introduce a tactile sensing system that employs a tactile estimation model trained on data generated in simulation. As a concrete implementation, we present the design and fabrication of a tactile balloon catheter based on an existing cardiac ablation balloon catheter, the HotBalloon catheter. We also describe the construction of a simulation system for collecting training datasets, as well as the architecture and training procedure of the tactile estimation model.

3.2 System configuration

The proposed data-driven balloon-catheter tactile sensing system consists of three main components: a data acquisition sequence, training sequence of Machine learning model, and an estimation and visualization sequence. The method aims to estimate tactile information from visual data, constructing a machine-learning model using a large amount of simulation-generated data, and subsequently inferring the three-dimensional shape of the balloon from optical marker patterns observed in real environments.

First, in the data acquisition sequence, the deformation behavior of the balloon catheter and the accompanying deformation of optical marker patterns are accurately simulated by combining deformable-body simulation with optical simulation. In the deformable-body simulation, material parameters of the balloon are specified, and deformation-shape data are generated to cover a wide variety of contact conditions. In the optical simulation, image data are obtained by reproducing how changes in the balloon surface geometry are projected into the image space of the endoscopic camera. Through this sequence of generation processes, a large-scale dataset is constructed that establishes correspondences between tactile information (deformation) and marker patterns.

Next, in the training sequence of Machine learning model, the aforementioned dataset is used to train a tactile estimation model that predicts the three-dimensional deformation shape of the balloon from optical marker patterns. In this study, a convolutional neural network (CNN) is adopted in order to capture both local and global deformation features of the marker patterns. The CNN extracts feature maps from image inputs and maps them, through a multi-layer learning pipeline, to three-dimensional shape information of the balloon. The model is trained in a supervised manner so that it learns the correspondence between input marker-pattern images and output three-dimensional meshes.

Finally, in the estimation and visualization sequence, marker-pattern images captured in real time by an endoscopic camera during catheter operation are analyzed, and the trained tactile estimation model is used to infer the balloon deformation shape. The inferred results are presented to the operator via three-dimensional visualization, enabling indirect understanding of the contact state of the balloon.

However, one of the most critical challenges in implementing this system is that the relative pose between the endoscopic camera and the balloon is not mechanically stable, and that there exists a Sim2Real gap between the deformable-body simulation and the real environment. In particular, uncertainty in camera pose can cause large variations in marker projections and thereby significantly degrade inference accuracy.

To address these issues, this study constructs a dataset in which domain randomization is applied during simulation-based data generation. Furthermore, as a preprocessing step in training, noise and geometric distortions are added to the simulation images to match the statistical properties of real data more closely, thereby improving the model’ s generalization ability with respect to uncertainties in camera pose. Through these measures, the system is designed to reduce inference errors arising from the Sim2Real gap and to enable robust tactile estimation under real-world conditions.

3.3 Tactile balloon catheter

In this study, we used the HotBalloon catheter developed by Toray Industries, Inc. , shown in Fig. 3.2a, as the base device for tactile sensor development. This device is a balloon catheter designed for the treatment of atrial fibrillation, in which a fluid inside the balloon is heated by RadioFrequency (RF) current and the resulting thermal energy is used to ablate myocardial tissue during pulmonary vein isolation [24, 25]. In the present work, we exploit the flexible, large-deformation balloon structure of this device as a skin for embedding vision-based tactile sensing functionality.

The appearance and design of the balloon part of the HotBalloon catheter are shown in Fig. 3.2a. The catheter body consists of a 5 mm-diameter outer shaft and a 2 mm-diameter inner shaft arranged along the central axis. The balloon is made primarily of polypropylene (PP), with an effective length of approximately 30 mm in the longitudinal direction. PP resin offers high flexibility, chemical stability, and heat resistance, and therefore possesses the properties required to withstand temperature increases and repeated deformation under ablation conditions. The proximal side of the balloon is bonded to the outer shaft and the distal side to the inner shaft using a photocurable resin. The balloon diameter varies between 5 and 35 mm depending on the volume of fluid injected into the interior. This large range of geometric change is designed to accommodate inter-patient anatomical variations in pulmonary vein morphology [24, 25].

In this section, we describe how vision-based tactile sensing functionality is added to the balloon structure.

3.3.1 Marker Pattern

The proposed method employs a marker-based approach. Therefore, multiple optical markers are regularly placed on the balloon surface to enable visual measurement of deformation. The markers serve as indicators that reflect local stretching, rotation, and curvature changes of the balloon surface, and are imaged in real time by an endoscopic camera. The resulting images are passed through camera calibration and geometric correction procedures (described later), and then provided as input to a machine-learning-based tactile estimation model. In this way, three-dimensional balloon surface geometry and external contact states can be estimated from visual changes in the marker patterns.

In this study, we investigated two types of optical marker patterns applied to the

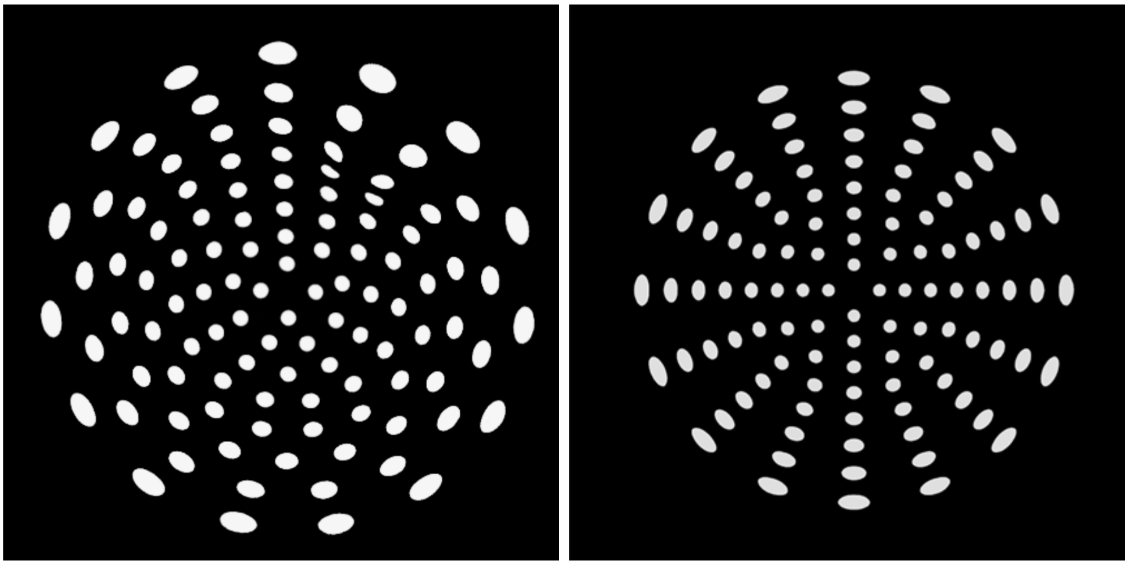

(a) (b)

Fig. 3.1: Marker layout patterns: (a) Asymmetric, (b) Symmetric

surface of the balloon catheter, shown in Fig. 3.1, based on the assumption that the marker layout strongly influences the accuracy and robustness of tactile estimation.

First, the asymmetric pattern shown in Fig. 3.1a was deliberately designed to avoid any rotational symmetry on the balloon surface. This asymmetry ensures that the geometric configuration of the marker set remains uniquely identifiable, even under small rotations or axial shifts of the endoscopic camera. Such uniqueness helps reduce ambiguities in depth and shape estimation. In vision-based tactile sensing, where only a single viewpoint is available, rotationally symmetric patterns can project similar visual appearances for different balloon orientations, potentially confusing the machine learning model. Thus, the asymmetric design offers improved discriminability for the estimation network.

In contrast, the symmetric pattern shown in Fig. 3.1b arranges the markers radially in a regular configuration. This layout facilitates a uniform visual marker density across the balloon surface, mitigating local information imbalance. If the marker spacing varies significantly, the tactile estimation model may exhibit heterogeneous reconstruction accuracy across regions; therefore, ensuring uniformity is important for stabilizing shape estimation.

We also examined the appropriate marker diameter. Markers with a diameter of 1 mm provide higher spatial resolution but frequently detached during fabrication due to insufficient adhesion area. Conversely, 3 mm markers were too large, reducing marker density and impairing the detection of fine local deformations necessary for tactile estimation. Considering both manufacturing robustness and sensing performance, we selected a marker diameter of 2 mm as an optimal balance.

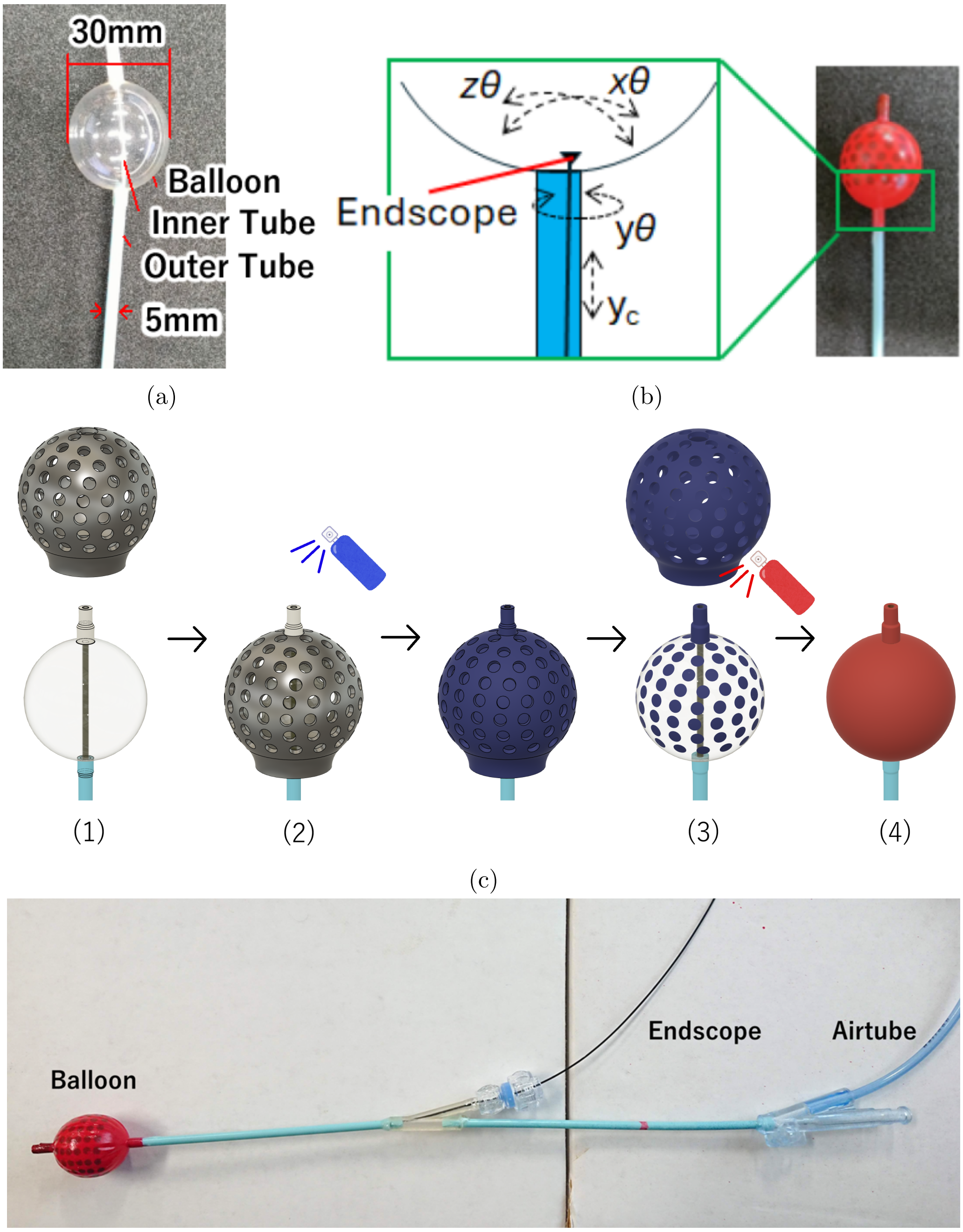
3.3.2 Fabrication

To apply optical markers to the balloon catheter surface with high precision and uniformity, we adopted a stencil-based coating process. An overview of the marker transfer procedure is shown in Fig. 3.2c.

First, the marker pattern to be transferred onto the balloon surface was geometrically optimized, and a corresponding stencil was designed in 3D CAD. The marker layout was designed so that markers are uniformly distributed over the entire circumference of the balloon. The stencil inner geometry was defined as a spherical shell whose inner length in the longitudinal direction matches the expanded balloon length of 30 mm. The stencil was fabricated using a fused deposition modeling (FDM) 3D printer. Since FDM printing is prone to edge rounding and dimensional inaccuracies due to layer-wise deposition, the marker diameter was set to 2 mm in the design, but the stencil apertures were printed as minimally sized 1 mm holes. The apertures were then mechanically enlarged using a pin vise as a post-processing step. This procedure removes staircase-like layer artifacts inherent to 3D printing and yields apertures with high circularity. As a result, the geometric accuracy of the coated markers is improved, reducing error factors in subsequent visual sensing and machine-learning-based tactile estimation.

Next, the balloon catheter is deflated and inserted into the stencil. The balloon is then inflated inside the stencil so that its surface comes into close contact with the stencil inner wall. Once contact is achieved, blue paint is sprayed through the stencil apertures to transfer the marker pattern onto the balloon surface. After coating, sufficient drying time is allowed to ensure that the paint adheres firmly.

After the blue paint has dried, the balloon is deflated again, removed from the stencil, and reinflated in free space. The entire balloon is then coated with red paint as an outer layer. Because the catheter itself is transparent, complex background patterns may be visible through the balloon if left uncoated. The red coating suppresses the influence of ambient light and background, thereby facilitating segmentation of the marker pattern. Red was chosen for the outer coating because it provides high optical contrast against



(a) (b)

(c)

(d)

Fig. 3.2: Design and fabrication of the TacBalloon catheter. (a) Design of the balloon catheter. (b) Design of the tactile balloon catheter and degrees of freedom of the endoscopic camera: the endoscope has lateral degrees of freedom of approximately 0. 5 mm, 3 mm in the axial direction (y_c), and up to 360 deg in rotation (y_θ). Due to bending of the balloon and catheter, the optical axis of the camera can tilt by up to approximately 45 deg relative to the catheter axis (x_θ and z_θ). (c) Fabrication of the tactile balloon catheter: (1) insert the balloon catheter into the 3D-printed stencil and inflate it inside; (2) spray blue paint through the stencil apertures; (3) remove the stencil after the paint has dried; (4) coat the balloon catheter with red paint. (d) Fabricated tactile balloon catheter

Table 3.1: Camera profile; Enable Inc. ENA-10448-AS

Diameter mm	Resolution px	FoV (air) deg	FoV (water) deg
1. 5	480×640	120	90

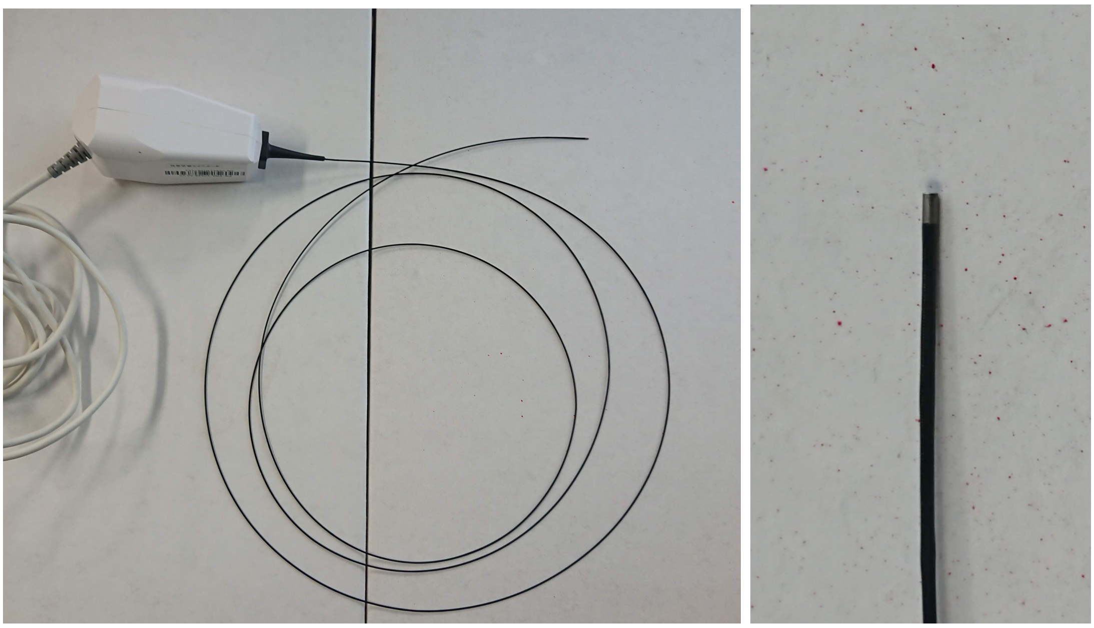
the blue markers and thus improves the accuracy of marker extraction by the endoscopic camera.

Since the balloon undergoes repeated inflation and deflation, conventional paints tend to crack easily, leading to degradation of marker shape. To avoid this problem, we used a polyurethane-based paint with high elasticity and excellent extensibility. Specifically, we used Urehero manufactured by Saito Paint Co. , Ltd. , ensuring homogeneous film thickness and mechanical flexibility.

3.3.3 Endoscopic camera

To enable imaging in the confined space inside the balloon catheter, we employed the ENA-10448-AS endoscopic camera manufactured by Enable Inc. , as shown in Fig. 3.3. This device has an outer diameter of 1. 5 mm, making it sufficiently small to be installed within the balloon catheter.

As illustrated in Fig. 3.2b, the endoscopic camera is inserted along the inner catheter



(a) (b)

Fig. 3.3: Endoscopic camera ENA-10448-AS. (a) Overall view of the endoscopic camera. (b) Close-up view of the distal tip.

toward the base of the balloon and positioned to observe the surface markers from inside the balloon. In the structure used in this study, the camera has no support structure near the balloon; it is mechanically constrained only by a stopcock-type sealing valve. Consequently, the camera has relatively large degrees of freedom: approximately 0. 5 mm in the lateral direction, 3 mm in the axial direction (y_c), and up to about 360 deg in rotation (y_θ). In addition, bending of the balloon and inner catheter can cause the optical axis of the endoscopic camera to tilt by up to approximately 45 deg relative to the central axis of the balloon catheter (x_θ and z_θ).

Furthermore, as shown in Figs. 3.4b and 3.4c, the inner shaft that maintains the balloon structure partially occludes the field of view of the endoscopic camera. As a result, a subset of the optical markers applied to the balloon surface is occluded and thus cannot be observed.

In the experiments, tactile estimation was performed with the balloon interior filled with either air or water. The field of view (FoV) of the endoscopic camera depends on the difference in refractive index between the camera lens and the surrounding medium. This relationship is described by Snell' s law in Eq. (3.1). When the balloon interior is filled with water ($n_2 \approx 1.333$), the effective FoV seen from the lens is reduced by approximately 33 % compared with the case where it is filled with air ($n_1 \approx 1.000$). Therefore, although the specified FoV of the ENA-10448-AS camera used in this study is 120 deg in air (Table 3.1), it is reduced to about 90 deg in water, as depicted in Fig. 3.4c. Consequently, the range of markers that can be simultaneously imaged is reduced.

Illumination for imaging inside the balloon catheter was provided by the LED light source mounted at the tip of the endoscopic camera.

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{3.1}$$

n_1 : refractive index of incident medium (air ≈ 1.000),

n_2 : refractive index of refracted medium (water ≈ 1.333),

θ_1 : incident angle, θ_2 : refracted angle

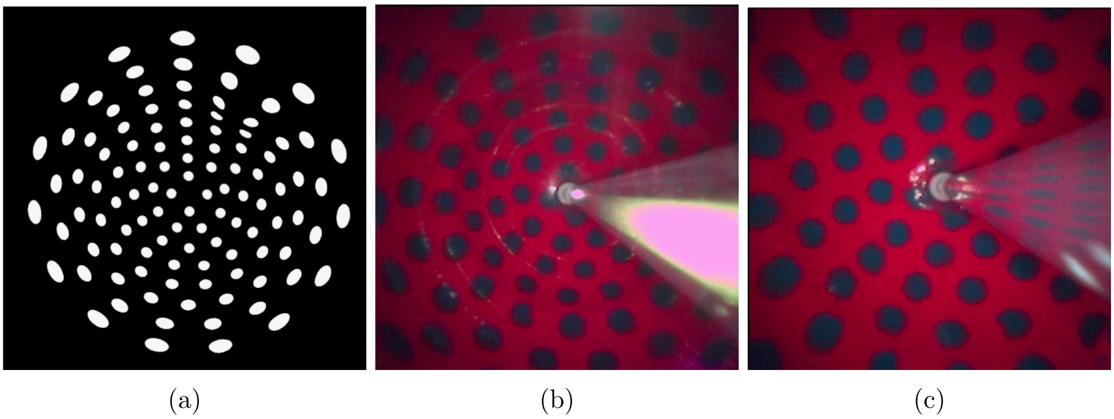

(a) (b) (c)

Fig. 3.4: Marker patterns and endoscopic camera views. (a) Reference marker layout (FoV = 180 deg). (b) Endoscopic image of the balloon catheter filled with air (FoV = 120 deg). (c) Endoscopic image of the balloon catheter filled with water (FoV = 90 deg).

3.4 Camera calibration

The optical system of the ENA-10448-AS endoscopic camera used in this study employs an ultra-wide-angle lens with a FoV of 120 deg. Compared with ordinary cameras, such ultra-wide-angle lenses produce images with pronounced barrel distortion. This optical distortion arises from the projection model of the lens. The standard projection model for conventional lenses is the pinhole camera model, expressed in Eq. (3.2), whereas ultra-wide-angle lenses are designed to follow projection models such as those in Eqs. (3.3)–(3.6).

$$y = f\frac{h}{x} \qquad \text{(pinhole model)} \tag{3.2}$$

$$y = f\theta \qquad \text{(equidistant projection)} \tag{3.3}$$

$$y = f\sin\theta \qquad \text{(orthographic projection)} \tag{3.4}$$

$$y = 2f\tan\left(\frac{\theta}{2}\right) \qquad \text{(stereographic projection)} \tag{3.5}$$

$$y = 2f\sin\left(\frac{\theta}{2}\right) \qquad \text{(equisolid-angle projection)} \tag{3.6}$$

y : image height, f : focal length, θ : angular coverage,

h : object height, x : distance to object

Optical lenses inevitably exhibit distortion due to manufacturing tolerances and subtle deviations in lens surface shape. Consequently, images obtained from the endoscopic camera do not strictly follow the idealized projection equations. Such optical distortions introduce significant geometric errors in marker positions and distances on the image plane. In particular, for the tactile estimation problem addressed in this study, small relative position changes between markers on the balloon surface are used to reconstruct the surface geometry. Without correction, systematic biases induced by distortion would be propagated into the reconstructed shapes, leading to a substantial degradation in estimation accuracy. Therefore, it is necessary to perform camera calibration to estimate the internal parameters of the endoscopic camera.

However, as mentioned above, the projection model of an ultra-wide-angle lens is not a simple pinhole camera model. Thus, calibration methods based on distortion models

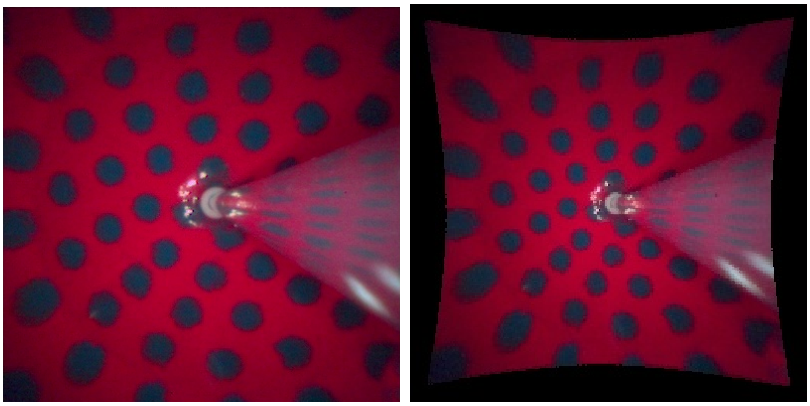

(a) (b)

Fig. 3.5: Distortion and correction of the ultra-wide-angle camera. (a) Original distorted image. (b) Image corrected by camera calibration.

derived from the pinhole model, such as the Brown–Conrady model [78], are not appropriate. Instead, we adopt the Ocam model proposed by Scaramuzza *et al.* as the distortion model [79]. The Ocam model is defined by a polynomial relationship as in Eq. (3.7) and is widely used as a standard method for fisheye camera calibration. It is also available as a fisheye camera calibration toolbox for MATLAB.

$$f = a_0 + a_1\rho^1 + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 \tag{3.7}$$

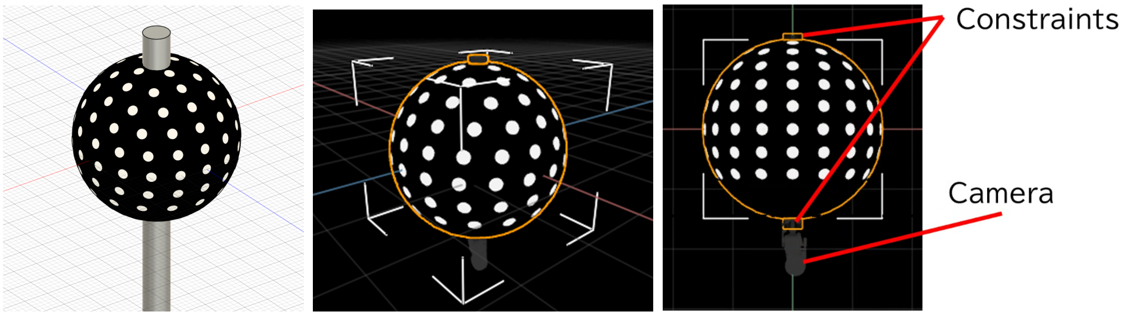
ρ : distance from the image center, a_n : distortion parameters

Using images of a checkerboard pattern captured from multiple poses, we estimated the parameters of the Ocam model. The estimated parameters include not only the polynomial coefficients a_n, but also the coordinates c of the distortion center. These parameters allow the fisheye images to be corrected into geometrically accurate images based on the equidistant projection model. Furthermore, separate calibrations were performed in air and in water. Because the optical path depends on the refractive index of the medium, a distinct optical model must be applied when the endoscope is used under water to achieve accurate distortion correction.

The estimated lens distortion parameters are listed in Table 3.2. The effect of distortion correction using these parameters is illustrated in Fig. 3.5.

Table 3.2: Distortion parameters

	a_0	a_1	a_2	a_3	a_4
air	111. 92	0	-2.06×10^{-4}	-1.09×10^{-5}	1.70×10^{-8}
water	168. 55	0	2.11×10^{-4}	-1.16×10^{-5}	1.07×10^{-8}



(a) (b) (c)

Fig. 3.6: Tactile Balloon Catheter and Simulation (a) 3D model of a tactile balloon catheter (b) Tactile balloon catheter constructed in simulation (c) Balloon catheter in simulation: constraints and virtual camera placement

3.5 Simulation

In this section, we describe the simulation framework used to reproduce the deformation behavior of the balloon catheter and the associated displacement of marker patterns, and to construct training datasets for the machine-learning model. The primary purpose of this simulation is to avoid the cost and experimental constraints of repeated trials on physical hardware, and to enable efficient acquisition of tactile data under a wide range of conditions for learning.

The simulation consists of two components: a deformable-body (soft-body) simulation that generates skin deformation caused by contact, and an optical simulation that generates endoscopic images of optical markers as they move with the deforming skin.

To integrate these two components, we built a simulation environment based on NVIDIA Isaac Sim [80]. Isaac Sim is a robotics simulation platform capable of executing high-fidelity physical simulation and optical rendering in real time. NVIDIA PhysX5 [81] is used for the physical simulation, in which the balloon deformation is analyzed as a soft body. This enables real-time computation of the balloon surface deformation under various conditions, including internal pressure variations, external contact forces, and kinematic constraints. In this study, we imported the 3D design model of the tactile balloon catheter fabricated in Section 3.3 into the simulation environment, modeled the balloon catheter and its optical marker patterns as a deformable body, and defined the projection model of the endoscopic camera to construct an integrated virtual experimental environment.

3.5.1 Soft-body simulation

The balloon catheter, which was designed in 3D CAD as illustrated in Fig. 3.6a, is placed in the simulation environment as shown in Fig. 3.6b. In particular, the balloon portion is modeled as a soft body. To reproduce the deformation behavior of the balloon catheter, we impose appropriate constraints on the balloon structure. Specifically, the mesh vertices at the proximal junction between the catheter shaft and the balloon are fixed in world coordinates and used as a reference frame for deformation.

The soft-body simulation is performed using NVIDIA PhysX5, which implements a finite-element-method (FEM)-based soft-body solver. In FEM-based soft-body simulation, the deformation behavior of the balloon can be reproduced by assigning material properties such as Young’ s modulus, Poisson’ s ratio, and density. Compared to mass–spring methods, which approximate deformable objects with point masses and springs, FEM provides better consistency with physical material parameters and higher accuracy in reproducing deformation shapes [82, 83]. However, even with FEM, accurately simulating the physics of highly deformable materials (e. g. , hyperelastic behavior) in real time remains challenging due to computational cost. Consequently, the soft-body simulation in PhysX5 does not perfectly match the true deformation behavior of the physical balloon.

Moreover, the balloon catheter treated in this study is a physically complex structure consisting of an elastic membrane and internal fluid. The apparent material properties of the balloon, such as effective Young’ s modulus, depend on deformation state, internal fluid pressure, and fluid viscosity. For this reason, it is difficult to reproduce the deformation behavior exactly with a single set of physical parameters.

In machine-learning-based applications that suffer from discrepancies between simulation and reality (the so-called Sim2Real gap), domain randomization—that is, randomizing physical parameters during dataset generation—has proven to be an effective strategy [84]. By training on data generated under a wide variety of physical conditions, the model can learn robust behavior and improve its generalization ability to real-world environments. In our setting, the dataset constructed by simulation consists of balloon deformation shapes and the corresponding endoscopic images; explicit force data applied to the balloon are not used. This makes it particularly suitable to adopt domain randomization of physical parameters.

For each contact episode in the simulation, we randomized the balloon’ s apparent

Table 3.3: Camera poses and simulation parameter ranges used for data collection. Min and Max denote the lower and upper bounds of each parameter range, respectively.

y (mm) Min — Max	y_θ (deg) Min — Max	x_θ (deg) Min — Max	z_θ (deg) Min — Max
-2 — 2	0 — 360	-40 — 40	-40 — 40

Young’ s modulus E (kPa) Min — Max	Coefficient of friction μ Min — Max	Poisson’ s ratio ν Min — Max
10 — 80	0 — 0. 1	0. 4 — 0. 5

Young’ s modulus E (kPa), Poisson’ s ratio ν, and coefficient of friction μ within the ranges listed in Table 3.3. This increases the diversity of deformation patterns experienced by the balloon and enables the construction of a robust dataset that covers the deformation shapes likely to occur in the physical device, thereby improving the estimation accuracy of the tactile model.

3.5.2 Optical simulation

The objective of the optical simulation is to generate virtual image data that correspond to actual endoscopic observations, by reproducing how deformation-induced changes in optical marker positions—obtained from the soft-body simulation—are captured by the endoscopic camera.

The endoscopic camera used in this study, ENA-10448-AS, is an ultra-wide-angle camera with a field of view of approximately 120 deg. While ultra-wide-angle cameras provide a wide field of view, they are prone to barrel distortion and complex nonlinear aberrations compared with standard lenses, and these distortions strongly depend on the adopted projection model. In this work, we adopt an equidistant projection model in which the distance y (in pixels) from the image center to any pixel is proportional to the incident angle θ (deg), as described by Eq. (3.3).

Within the simulation environment, we construct a virtual camera that follows the equidistant projection model of Eq. (3.3). The field of view of the virtual camera is set to 180 deg, which is wider than that of the actual endoscopic camera. This choice ensures that all optical markers on the balloon surface are within the imaging range when generating training datasets. The discrepancy between the virtual camera’ s field of view and that of the real camera is handled later during the machine-learning preprocessing stage. Because, under the equidistant projection model, the distance y (pixels) from the image center is proportional to the incident angle θ, we simulate the limited field of view of the real camera by cropping the simulation images to the desired angular range during preprocessing.

For optical rendering, we use the rendering engine provided by NVIDIA Isaac Sim and treat the camera pose as a variable. Since images are binarized before entering the machine-learning pipeline, we configure the optical simulation to output images in which markers are rendered in white and all other regions in black, thereby generating images that contain only the marker shapes (see e. g. , Fig. 3.5a).

3.5.3 Collecting training data

To obtain the data required for training the machine-learning model, we apply diverse contact interactions to the balloon catheter within the simulation environment and construct a dataset consisting of balloon deformation behaviors and the corresponding endoscopic images. The dataset is designed to cover a wide range of contact conditions that may occur in practice, thereby improving the generalization performance of the model. At each simulation step, we collect the following two types of data for learning:

- **Vertex coordinates of the deformed balloon mesh:** The three-dimensional coordinates (x, y, z) of 642 vertices of the deformed balloon mesh, as computed by the soft-body simulation.
- **Virtual endoscopic camera images:** Images captured by the virtual camera that depict the deformation and displacement of optical markers induced by the balloon deformation in the soft-body simulation.

As contacting objects in the simulation, we use three geometric primitives: a sphere, a cylinder, and a rectangular block. For each simulation episode, we randomize the number of contact objects, the aspect ratios of their shapes, and their contact directions, positions, and indentation depths, in order to generate a wide variety of deformation patterns. In addition, the pose of the virtual endoscopic camera—specifically, its position along the optical axis y_c and its tilt angles x_θ and z_θ (Fig. 3.2b)—is randomized within the ranges listed in Table 3.3. This allows us to collect image data that reflect field-of-view changes caused by camera misalignment and position errors, thereby bringing the simulated images closer to those observed in real environments.

Furthermore, to reduce the Sim2Real gap, we apply domain randomization to the physical parameters used in the soft-body simulation, namely Young’ s modulus E, the static coefficient of friction μ, and Poisson’ s ratio ν. The randomization ranges are shown in Table 3.3, where “Under” and “Max” denote the lower and upper bounds, respectively; parameter values are sampled uniformly within these ranges for each episode.

In total, approximately 70, 000 pairs of contact patterns and endoscopic images were collected, and a large-scale dataset was constructed for training the tactile estimation model.

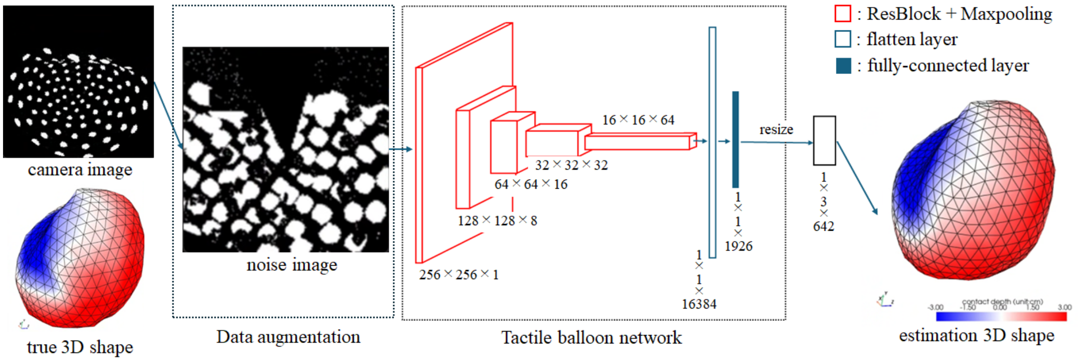


Fig. 3.7: Tactile Balloon Network: simulation endoscopic images are used as inputs, and the corresponding deformed balloon 3D meshes are used as ground-truth labels for supervised learning. By training the network with noise-augmented endoscopic images, the model learns to generalize and becomes capable of estimation from real endoscopic images.

3.6 Machine Learning

Using the simulation dataset constructed in the environment described in Section 3.5—which consists of endoscopic images and the corresponding deformed balloon meshes—we construct a tactile estimation model, referred to as the *Tactile Balloon Network*, and train it in a supervised manner to estimate tactile information. In addition, to account for noise components that cannot be fully represented in simulation, we apply data augmentation during training so that the resulting model acquires robustness to various types of noise encountered in real environments.

3.6.1 Preprocessing of training data

As described in Section 3.5, the endoscopic images used as inputs to the machine learning model are idealized images generated by a virtual camera in the simulation environment. These images do not contain the various disturbances present in actual surgical settings, such as lens aberrations, illumination variations, scattering in fluid, marker defects, and reflections, and thus constitute a noise-free dataset. By contrast, real endoscopic images are affected by diverse optical, mechanical, and environmental noise sources, as summarized in Table 3.4. A model trained solely on such clean simulation images would be vulnerable to these disturbances, and its estimation accuracy would inevitably degrade due to the Sim2Real gap.

Reproducing all such real-world noise phenomena in simulation is difficult. Moreover, multiple noise sources often occur simultaneously, making it impractical to generate a dataset that exhaustively covers all combinations. Therefore, in this study we collect only noise-free camera images under idealized conditions at the data generation stage, and subsequently enhance adaptability to real environments by applying data augmentation as a preprocessing step during training.

Concretely, as illustrated in Fig. 3.7, we adopt a strategy in which random noise is generated and added with a certain probability to each simulation image immediately before it is fed to the network. The noise types used are listed in Table 3.4 and include brightness variation, Gaussian noise, random masking, image rotation or translation , cropping , and blur, thereby mimicking various real-world disturbances. Because the combination of these noise sources is determined stochastically, even the same input image is subjected to different noise patterns at each epoch, as illustrated in Fig. 3.12. This design encourages the model to acquire high robustness against real-world noise.

The original dataset consists of approximately 70, 000 noise-free images obtained in simulation. At each epoch, however, new random noise is applied to the images during preprocessing, so that the same sample is seen under different noise conditions throughout training. To exploit this effect, we set the total amount of training iterations to be 100 times the nominal dataset size. In this way, we can keep the size of the underlying dataset modest while still training a tactile estimation model that minimizes the gap between simulation images and real images.

Finally, after adding noise, we apply binarization and resizing to the images and reshape them into $256 \times 256 \times 1$ tensors, which constitute the input format of the Tactile Balloon Network.

3.6.2 Machine learning model

The overall structure of the training process is illustrated in Fig. 3.8. The Tactile Balloon Network developed in this study takes a preprocessed endoscopic image ($256 \times 256 \times 1$) as input and is formulated as a multivariate regression model that directly estimates the balloon’ s three-dimensional shape.

The network consists of five Residual Blocks [85], with max-pooling layers inserted between blocks to progressively reduce spatial resolution. This design enables extraction of both global and local features, capturing fine marker displacements on the balloon surface as well as overall deformation trends. After feature extraction through the five residual

Table 3.4: Noise Models and Data Augmentation

Data augmentation	Cropping	Translation
Noise in data	Random cropping of image	Random translation of image
Real-world noise	Camera motion Difference in field of view due to medium	Camera motion
Variables	Crop ratio: s [%] Crop origin: (sx_0, sy_0)	Translation offset: (tx, ty)[pixel]
Range	$40 \leq s \leq 80$ $0 \leq sx_0 \leq 256(1-s)$ $0 \leq sy_0 \leq 256(1-s)$	$-50 \leq tx \leq 50$ $-50 \leq ty \leq 50$

Data augmentation	Rotation	Brightness / Contrast
Noise in data	Random rotation of image	Random gamma correction and brightness variation
Real-world noise	Camera rotation	Illumination variation
Variables	Rotation angle: y_θ [deg]	pixel value: $c' = c \times \gamma + \beta$ Gamma scale: γ Brightness: β
Range	$0 \leq y_\theta \leq 360$	$0.8 \leq \alpha \leq 1.2$ $-20 \leq \beta \leq 20$

Data augmentation	Gaussian noise	Region masking
Noise in data	Addition of Gaussian distributed noise	Random masking of image
Real-world noise	Turbidity in underwater environments	Marker loss caused by specular reflection and halation wrinkles on the balloon surface occlusion by the inner tube
Variables	Mean of Gaussian: $mean$ Standard deviation: σ	Mask size: (mlx, mly)[pixel] Mask origin: (mlx_0, mly_0)[pixel] Masking ratio: m[%]
Range	$0 \leq mean \leq 0$ $0 \leq \sigma \leq 0.002$	$10 \leq lx \leq 128$ $10 \leq ly \leq 128$ $0 \leq lx_0 \leq 256$ $0 \leq ly_0 \leq 256$ $1 \leq m \leq 40$

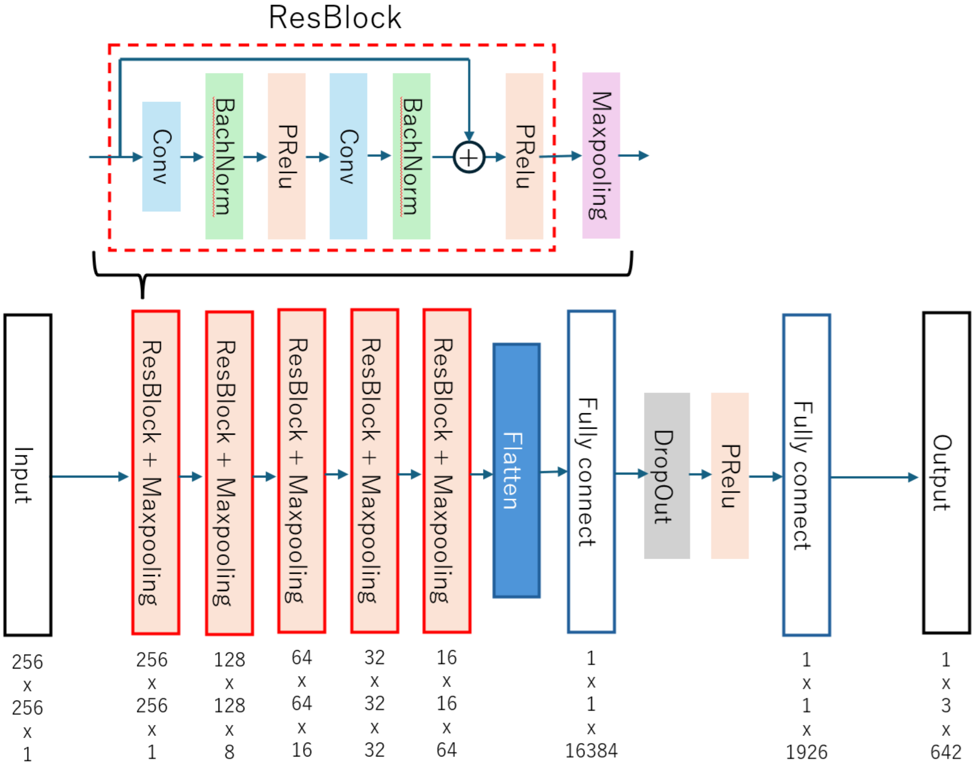


Fig. 3.8: Architecture of the tactile estimation network: a 256 × 256 × 1 image is fed as input. Five stages of residual blocks and max-pooling layers are applied. The resulting tensor is flattened and processed by fully connected layers with dropout and activation functions. Finally, the network outputs the 3D coordinates of the balloon mesh vertices.

blocks, the resulting feature maps are flattened and passed through fully connected layers to generate a 1926-dimensional output corresponding to the x, y, and z coordinates of the 642 vertices forming the balloon’ s 3D mesh.

PReLU [86] is used as the activation function in each residual block. To mitigate overfitting, dropout (30%) is introduced between the fully connected layers. This is intended to improve generalization during training and enhance robustness to inputs that include real-world noise.

Note that we do not employ recurrent architectures such as LSTMs that utilize temporal information. While time-series models could potentially improve shape estimation accuracy by referencing adjacent frames, their use would require collecting training data that comprehensively includes continuous deformation histories leading to each deformation state, resulting in an excessively large dataset. In addition, temporal variations in camera pose and optical noise would also need to be modeled. Given the difficulty of preparing such sequential data, we adopt a CNN-based framework with single-image input and perform tactile estimation independently for each frame, prioritizing both data collection efficiency and robustness in real-world settings.

Training is performed in a supervised manner using the 3D balloon models from the simulation dataset as ground-truth labels. As the loss function, we adopt the mean squared error (MSE) defined in Eq. (3.8). The MSE evaluates the Euclidean distance between the predicted and ground-truth coordinates for each mesh vertex and thus directly penalizes shape reconstruction error.

$$\begin{aligned} L_{\mathrm{MSE}} &= \frac{1}{n}\sum_{i=1}^{n}\left\| \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} - \begin{pmatrix} x_i' \\ y_i' \\ z_i' \end{pmatrix} \right\|^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2\right], \end{aligned} \tag{3.8}$$

where (x_i, y_i, z_i) and (x_i', y_i', z_i') denote the predicted and ground-truth coordinates of the i-th mesh vertex, respectively, and n is the total number of vertices.

For optimization, we use stochastic gradient descent (SGD) with an initial learning rate of 1×10^{-4}. Additionally, if no improvement in the loss is observed over five consecutive epochs, the learning rate is reduced by a factor of 10, thereby preventing training stagnation and premature convergence to suboptimal local minima.

With this configuration, the Tactile Balloon Network operates as a deep-learning model that can accurately estimate the three-dimensional deformation of the balloon from visual information alone, providing a foundation for real-time reconstruction of contact geometry from endoscopic images.

3.6.3 Training results

The training conditions were as follows: the initial learning rate was set to 1×10^{-4}, and if no improvement in the loss was observed over five consecutive epochs, the learning rate was reduced by a factor of 1/10. This schedule encourages broad exploration in the early training phase and finer error minimization as training progresses, leading to stable convergence. As in the previous subsection, we used SGD as the optimizer and the MSE defined in Eq. (3.8) as the loss function.

Although the dataset comprises 70, 000 noise-free simulation images, random noise is applied at each epoch during preprocessing, so that from the perspective of the model, it effectively receives data with a diversity equivalent to roughly 100 times the dataset size. This design enables stable estimation even in the presence of disturbances associated with the Sim2Real gap. The batch size is set to 128, and the dataset is split into training and validation subsets with a ratio of 8:2 for cross-validation.

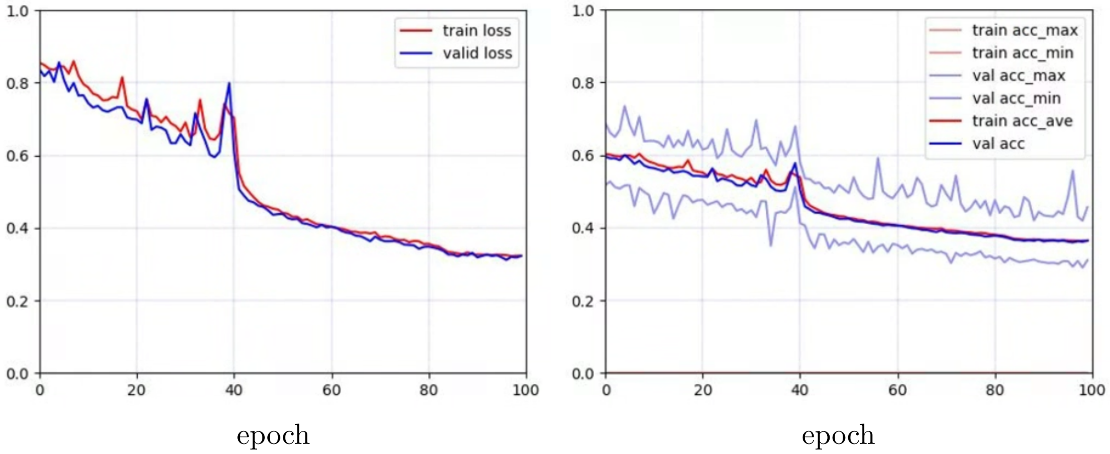
Training was conducted on a workstation equipped with an Intel Core i9 processor and an NVIDIA GeForce RTX 4080 GPU (16 GB VRAM). The high degree of parallelism offered by the GPU, accessed via CUDA, enabled efficient processing of the large-scale training with extensive data augmentation. The learning system and network model were implemented in PyTorch. Training was performed for 100 epochs, with the total training time being approximately 23. 5 hours.

The training curves for the asymmetric marker pattern shown in Fig. 3.1a are presented in Fig. 3.9. Both curves decrease steadily as the number of epochs increases; the learning rate is reduced around 40 epochs, after which the error continues to decrease more gradually and reaches a convergent trend around 80 epochs. This behavior indicates that the learning rate scheduling functioned appropriately, enabling stable optimization while suppressing overfitting.

In contrast, for the symmetric marker pattern, the loss did not decrease sufficiently. Figure 3.10 shows the training curves obtained when the upper bound of the Rotation augmentation, which mimics the camera rotation angle y_θ, was set to 45 deg, i. e. , less than half of the symmetric angle of 90 deg of the marker layout. For comparison, Fig. 3.11 shows the training curves for the asymmetric marker pattern under the same rotational augmentation condition. While the asymmetric pattern enables stable learning regardless of the magnitude of the rotation angle, the symmetric pattern yields stable training only when the allowable rotation range is relatively small.

From these observations, we conclude that the asymmetric marker pattern shown in Fig. 3.1a is more suitable for tactile sensors such as TacBalloon, in which the camera has rotational degrees of freedom as illustrated in Fig. 3.2d. Therefore, in the experiments described in expriment, we adopt the asymmetric marker pattern.

Figure 3.12 shows examples of input images and their corresponding 3D balloon meshes in the training dataset, as well as noise-augmented images fed into the Tactile Balloon

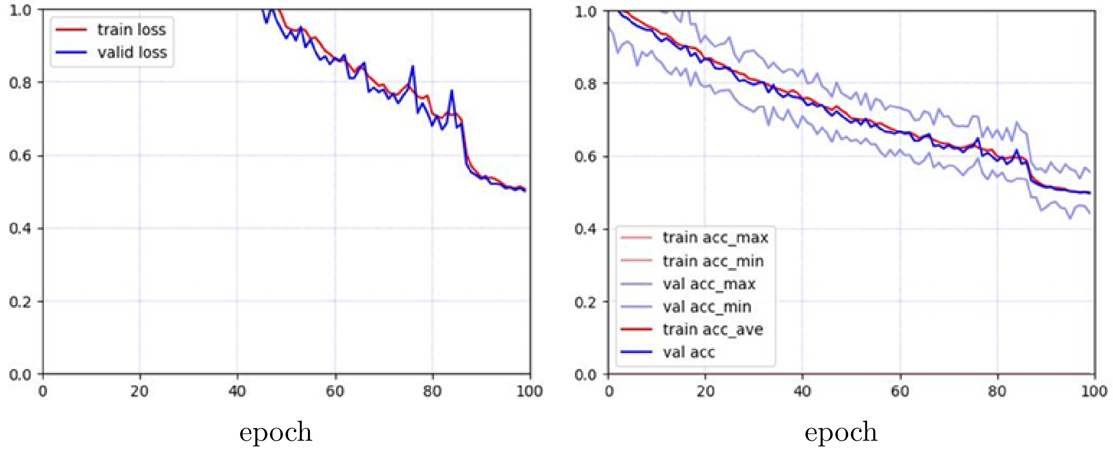


(a) (b)

Fig. 3.9: Training curves for the asymmetric marker pattern. Smaller values indicate higher estimation accuracy of the model. (a) Loss function (b) Least mean square error.

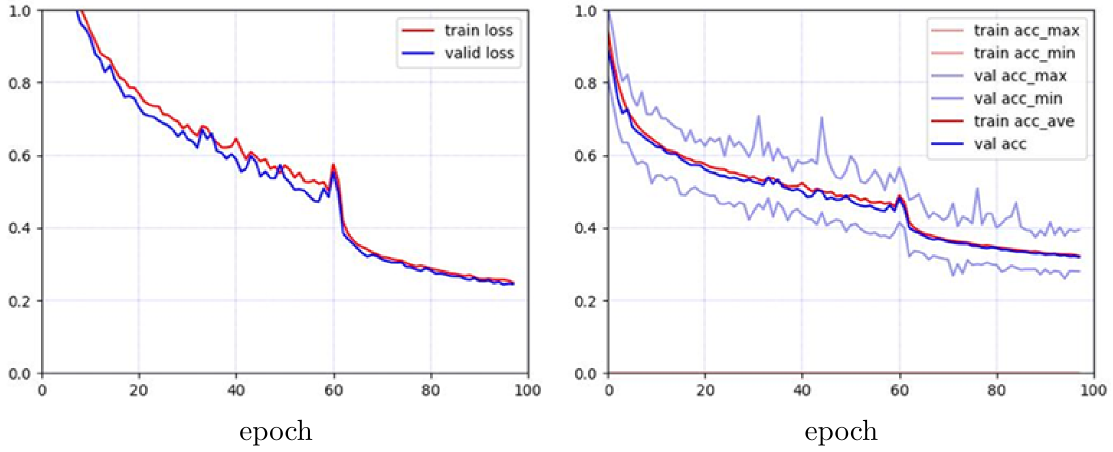
Network and the resulting predicted meshes. The color of each 3D mesh encodes the depth of deformation: relative to the undeformed 30 mm sphere, vertices that deform inward toward the center are shown in blue, whereas those that deform outward are shown in red. The predicted results faithfully reproduce features such as balloon inflation, indentation, and localized deformation under varying contact directions, depths, and noise patterns. Notably, even in regions near the boundary of the field of view and for complex deformation patterns, the predicted meshes do not exhibit severe artifacts, suggesting that the proposed method has strong representational capacity for diverse contact conditions.

Using the trained model for real-world inference, the system executes a full processing pipeline consisting of camera input acquisition, calibration, preprocessing and tensor formatting, model inference, and output of the reconstructed 3D shape data. When running on an Intel Core i7 processor with an NVIDIA GeForce RTX 3060 GPU (12 GB VRAM), the inference operated at approximately 24 fps, while CPU-only execution on an AMD Ryzen 5 6600U achieved around 10 fps.



(a) (b)

Fig. 3.10: Training curves for the symmetric marker pattern with a small upper bound on the rotation angle. Smaller values indicate higher estimation accuracy of the model. (a) Loss function (b) Least mean square error.



(a) (b)

Fig. 3.11: Training curves for the asymmetric marker pattern with a small upper bound on the rotation angle. Smaller values indicate higher estimation accuracy of the model. (a) Loss function (b) Least mean square error.

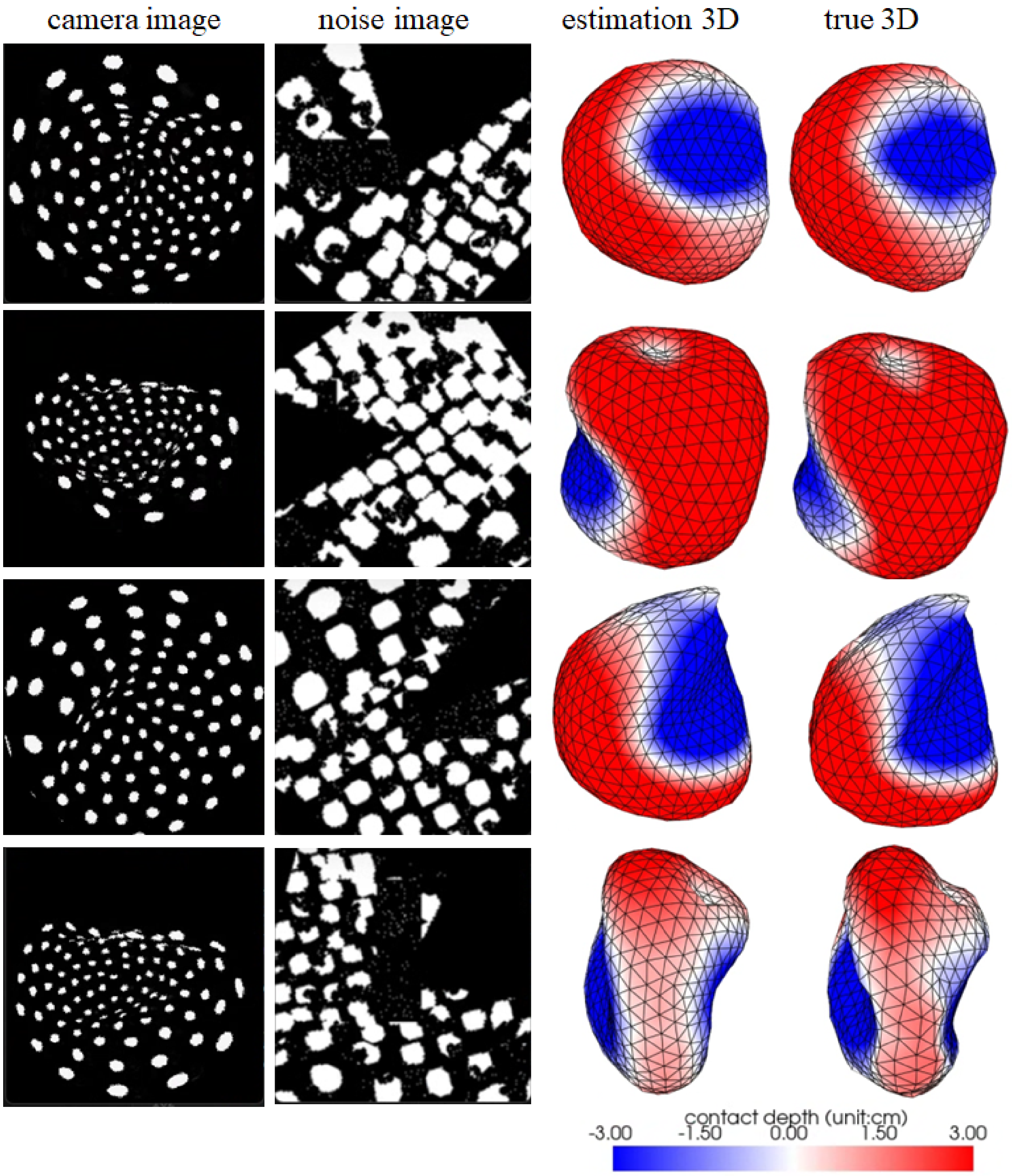


Fig. 3.12: Inference results: from left to right, raw endoscopic image, preprocessed input image, predicted balloon shape, and ground-truth mesh. The network demonstrates stable shape estimation performance even when noise is added.

Chapter4

Experiments and Performance Evaluation

4.1 Overview

This chapter describes a series of experiments conducted to comprehensively evaluate the performance of the proposed information-driven tactile balloon catheter system and its core component, the Tactile Balloon Network. The objective of these experiments is to quantitatively verify whether the model, which has learned tactile information in a simulation environment, can be successfully applied to real physical situations to achieve accurate deformation estimation, depth estimation, and contact shape reconstruction. To assess the effectiveness of the proposed model, we designed three groups of experiments from the following perspectives:

- Depth estimation experiment
  We evaluate the estimation accuracy of indentation depth (contact displacement) when the balloon comes into contact with an external object. By using test structures with known indentation depths and comparing the estimated results with the ground truth, we verify the point-wise deformation reconstruction performance of the tactile model.

- Shape reconstruction experiment
  We evaluate the three-dimensional shape estimation accuracy of the model under conditions where the overall balloon surface is constrained by a known reference geometry. By inflating the balloon inside a hollow spherical structure and computing the error between the estimated three-dimensional balloon shape and the known spherical surface, we clarify the global shape reconstruction performance and the spatial distribution of reconstruction accuracy.

- Contact area estimation experiment

  We bring the balloon catheter into contact with models of actual anatomical structures, such as cardiac tissue, and evaluate the estimation accuracy of the area and shape of the contact region. The ground truth is obtained from contact regions reconstructed by photogrammetry, which are compared with the contact regions estimated by the tactile model.

4.2 Depth Estimation Accuracy

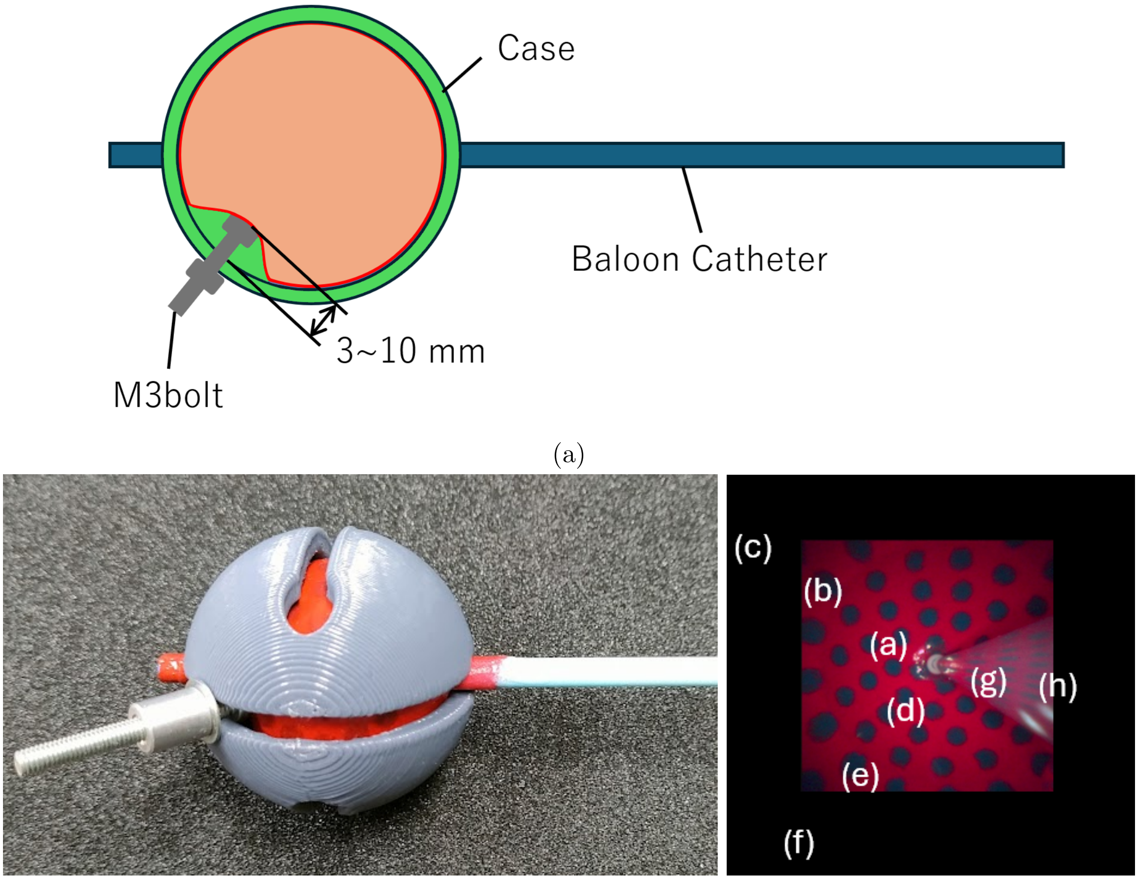
To quantitatively evaluate the estimation accuracy of the trained Tactile Balloon Network, we constructed an experimental system with a known geometric configuration and contact depth. In this experiment, we focus on the estimation of deformation depth as a representative point-wise evaluation of deformation estimation, and investigate the depth estimation performance under various conditions. In addition to indentation depth, we evaluate how the estimation accuracy is affected by the contact position, particularly whether the contact lies inside or outside the camera field of view and whether the optical markers on the balloon surface are partially occluded by internal structures.

4.2.1 Experimental Design

The depth estimation experiment was conducted using an evaluation system with the structure shown in Fig. 4.1a. In this experiment, the tactile balloon catheter is first inserted into a hollow spherical case and inflated. An M3 bolt mounted as the indenter is then advanced from the case wall to impose a prescribed indentation on the balloon surface, and the estimation accuracy of the corresponding deformation depth is evaluated. The actual experimental setup is shown in Fig. 4.1b. The case structure was fabricated from polylactic acid (PLA) using a fused deposition modeling (FDM) 3D printer. The inner diameter of the case is 30 mm, and it is equipped with an insertion port to hold the balloon catheter in place. In addition, a linear guide groove is formed on the outer surface of the case, along which the M3 bolt can be translated. This allows the contact position and indentation depth to be stably controlled.

The contact positions were defined based on the field of view of the endoscopic camera, and eight points (a–h) were placed as shown in Fig. 4.1c. These positions include points lying within the field of view (a, b, d, e, g), points outside the field of view (c, f, h), and points that are partially occluded by the inner tube (g, h). Furthermore, the presence and density of markers were taken into account: points a and b were placed on marker locations, whereas points d and e were placed between markers, enabling evaluation of how marker density affects estimation performance.

The indentation depth was defined relative to the state in which the balloon just comes into contact with the inner spherical wall, which was taken as the baseline (0 mm). The position of the bolt head was then adjusted to vary the indentation depth from 3 mm to 10



(a)

(b)

(c)

Fig. 4.1: Experimental design of the depth evaluation experiment. (a) Schematic of the experiment: the tactile balloon catheter is inserted into the inner cavity of the case and inflated. An M3 bolt mounted on the case is advanced to impose a prescribed indentation depth on the balloon surface. (b) Photograph of the actual experimental setup. (c) Contact positions: each contact point has the following characteristics: (a, b, d, e, g) positions within the camera field of view; (c, f, h) positions outside the field of view; (g, h) positions that are partially occluded by the inner tube.

mm in 1 mm increments. This design covers a wide range of deformation conditions, from small to relatively large deformations. The insertion axis of the bolt was kept constant to avoid directional variability at a given depth.

For each combination of contact position and indentation depth, measurements were repeated 10 times. The average and variance of the estimated indentation depth and deformation were then computed to assess the stability and accuracy of the estimation.

4.2.2 Experimental Results

The depth estimation accuracy as a function of deformation magnitude and contact position is summarized in Fig. 4.2. For contact points located within the camera field of view and not occluded by internal structures (a, b, d, e), the mean estimation error remained within approximately $\pm 1\,mm$ across the entire depth range from 3 mm to 10 mm, indicating high estimation accuracy. This result suggests that, when marker patterns are clearly captured and local deformation patterns are sufficiently provided to the network, the three-dimensional shape of the balloon can be stably reconstructed.

By contrast, at contact points located outside the field of view (c, f) and at positions partially occluded by the inner tube (g, h), the mean error increased to approximately $\pm 2\text{–}3\,mm$. In particular, at the occluded positions, the disappearance of markers from the captured image led to noticeably larger estimation errors. Moreover, even at in-view locations, the error tended to increase for larger indentation depths; slight underestimation of the indentation was observed around 10 mm.

4.2.3 Discussion

Table 4.1 summarizes the results of F-tests conducted between each contact point, using the error distributions aggregated over the indentation depths from 1 mm to 10 mm.

In the unoccluded regions within the endoscopic field of view (a, b, d, e), an accuracy within ±1 mm was achieved. This indicates that the Tactile Balloon Network can reliably recognize marker displacements and accurately reconstruct the three-dimensional deformation even on the physical device. This result suggests that the noise augmentation strategy and deep network architecture introduced in this study function robustly against various noise conditions that are close to those encountered in real environments.

In contrast, in the out-of-view regions (c, f, g, h), the errors show statistically significant differences compared to the in-view regions ($p < 0.05$). The primary cause of this increased error is the loss of input information due to the absence of visible marker information in the image. In out-of-view regions, local deformation features cannot be directly obtained, and therefore the network must infer indentation depth indirectly from global shape changes appearing in neighboring areas, which leads to reduced estimation accuracy. Furthermore, in the regions occluded by the inner shaft (g, h), markers are partially missing even within the nominal field of view, which likely causes the loss of local features necessary for accurate

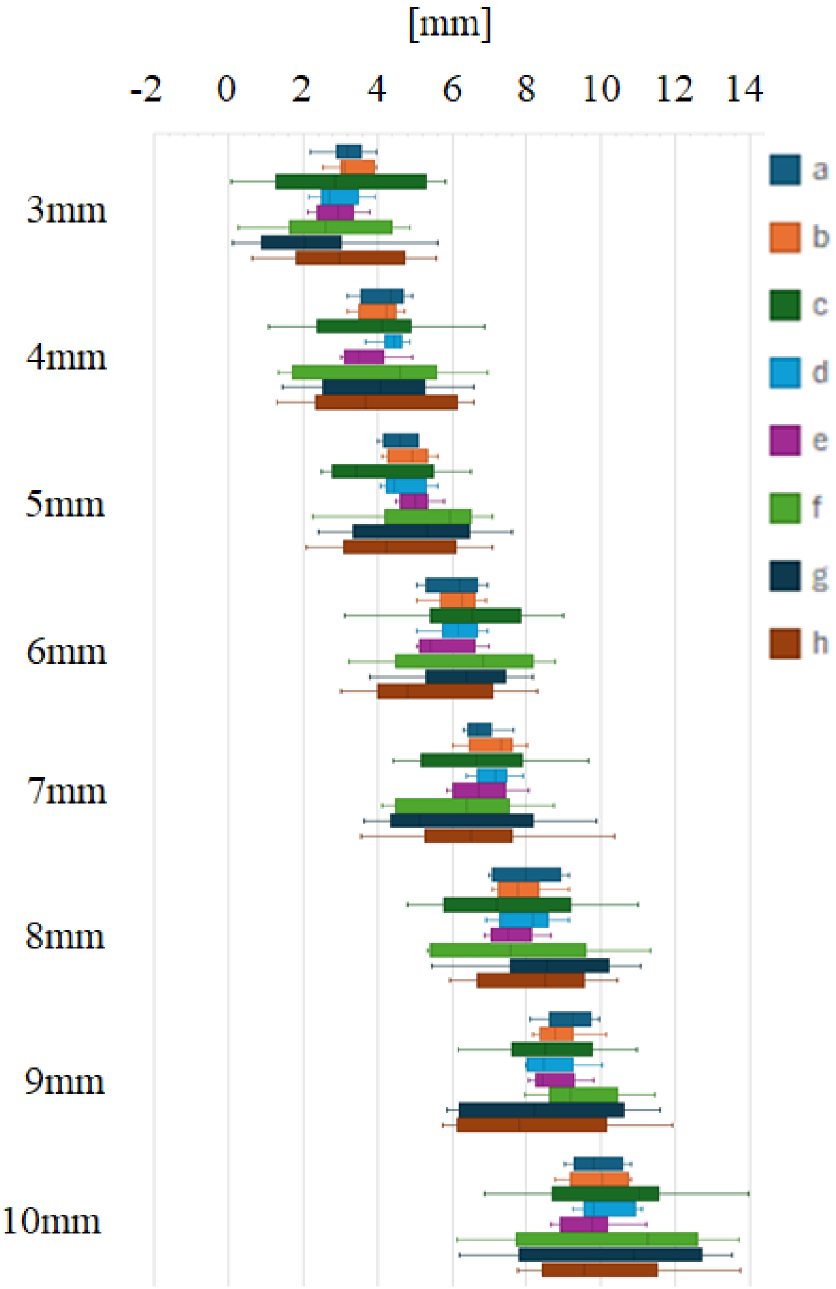


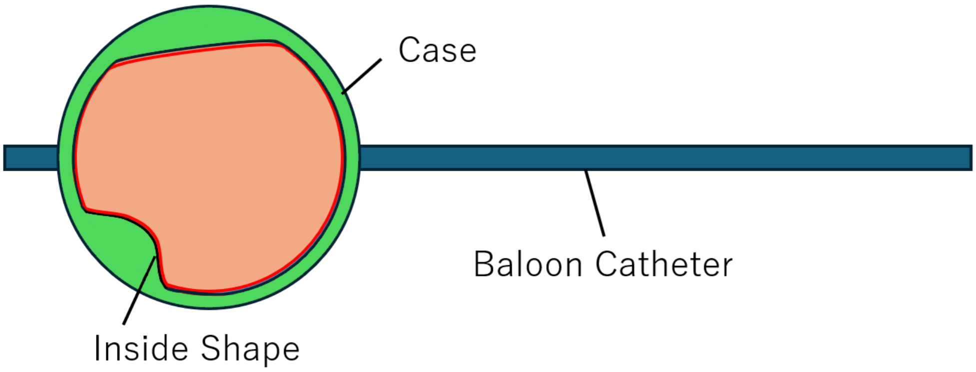
Fig. 4.2: Results of the depth estimation experiment. At positions a, b, d, and e, where markers are directly visible, depth estimation exhibits stable and consistent accuracy. In contrast, at positions c, f, g, and h, where markers lie outside the field of view or are occluded, the estimation accuracy significantly deteriorates.

Table 4.1: Results of F-tests between contact points

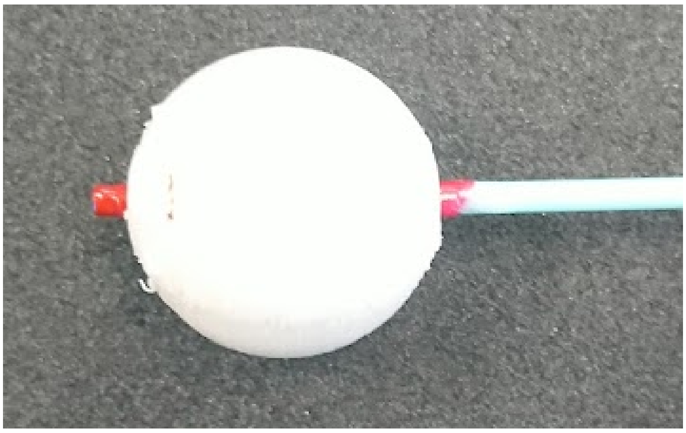
	b	c	d	e	f	g	h
a	0.712	0.000	0.958	0.825	0.000	0.000	0.000
b	-	0.000	0.752	0.556	0.000	0.000	0.000
c	-	-	0.000	0.000	0.598	0.575	0.758
d	-	-	-	0.785	0.000	0.000	0.000
e	-	-	-	-	0.000	0.000	0.000
f	-	-	-	-	-	0.973	0.826
g	-	-	-	-	-	-	0.800

shape reconstruction and thus contributes to the increased error.

On the other hand, no statistically significant differences were observed among the errors at the contact points within the visible region ($p > 0.05$). This suggests that the estimation accuracy does not strongly depend on the specific (asymmetric) marker pattern within the field of view.



(a)



(b)

Fig. 4.3: Experimental design of the shape estimation experiment. (a) Schematic of the experiment: the tactile balloon catheter is inserted into the inner cavity of the case and inflated so that the balloon follows the internal shape of the case. (b) Photograph of the actual setup used for the shape estimation experiment.

4.3 Surface Deformation Accuracy

To further quantify the estimation accuracy of the trained Tactile Balloon Network, we constructed an experimental system with a known global geometry. In this experiment, we evaluate the shape estimation accuracy of the balloon from a surface-wise perspective, focusing on the reconstruction accuracy of the overall deformed geometry.

4.3.1 Experimental Design

The shape estimation experiment was performed using the evaluation system shown in Fig. 4.3a. In this setup, the balloon catheter is inserted into a case with a hollow internal structure and inflated at a prescribed pressure such that the balloon deforms along the inner wall geometry. Because the internal wall shape of the case is designed as a known reference surface, the balloon surface passively conforms to this geometry, allowing quantitative evaluation of the balloon surface mesh estimated by the Tactile Balloon Network.

The actual experimental setup is shown in Fig. 4.3b. The case was fabricated from PLA using an FDM 3D printer. The internal cavity is based on a sphere of 30 mm in diameter, onto which four distinct deformation patterns were designed, as illustrated in Fig. 4.4b. These patterns incorporate geometric variations such as indentations, protrusions, and asymmetries to evaluate the estimation performance under diverse surface deformation conditions. An insertion port is provided on the side of the case to stably hold the balloon catheter at a prescribed position.

We further evaluate the deformation shape estimation performance by extracting the local geometry corresponding to the region where the balloon catheter is placed in contact, based on the two pulmonary-vein shape which mesh obtained from the NIH heart library [87] shown in Fig. 4.5.

In this evaluation, the balloon surface is passively driven to conform to a known shape, and the ability of the Tactile Balloon Network to reconstruct the three-dimensional balloon geometry solely from visual information is examined. In contrast to the depth estimation experiment, which focuses on local deformation, this experiment targets the global shape reconstruction capability of the model.

4.3.2 Experimental Results

Figure 4.4a shows the errors between all vertices of the estimated three-dimensional balloon shape and the corresponding vertices of the ground-truth shapes. For all four target shapes (Shape 1–4), the estimation error converged to approximately $\pm 1\,\mathrm{mm}$, indicating high shape reconstruction accuracy overall.

Figure 4.4b visualizes the distance between each estimated vertex and its ground-truth counterpart for each target shape. White regions indicate low error, while colors with increasing intensity represent larger errors. Red regions indicate vertices that are estimated outside the ground-truth surface, whereas blue regions represent vertices estimated inside the ground-truth surface. The arrow in each panel points toward the base side of the balloon catheter, which lies outside the field of view of the endoscopic camera.

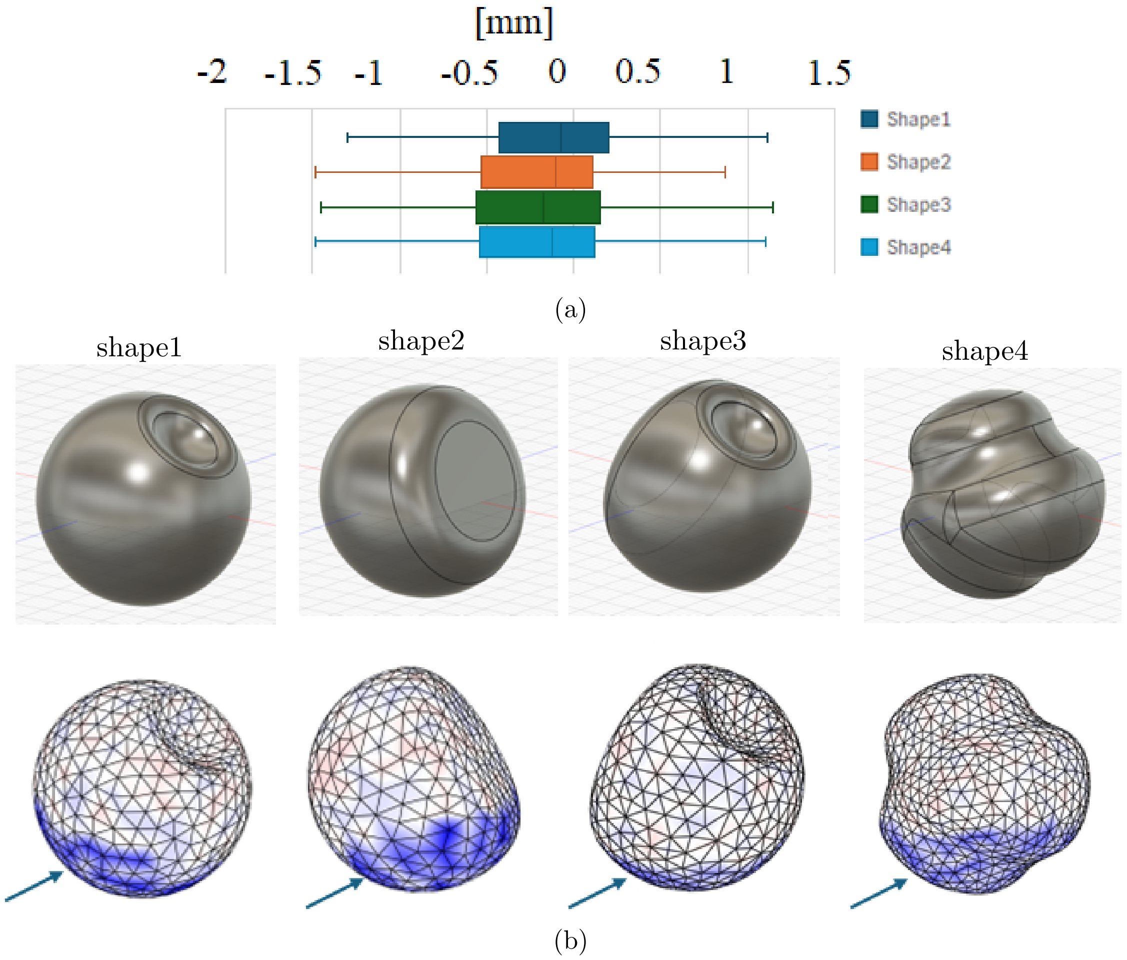
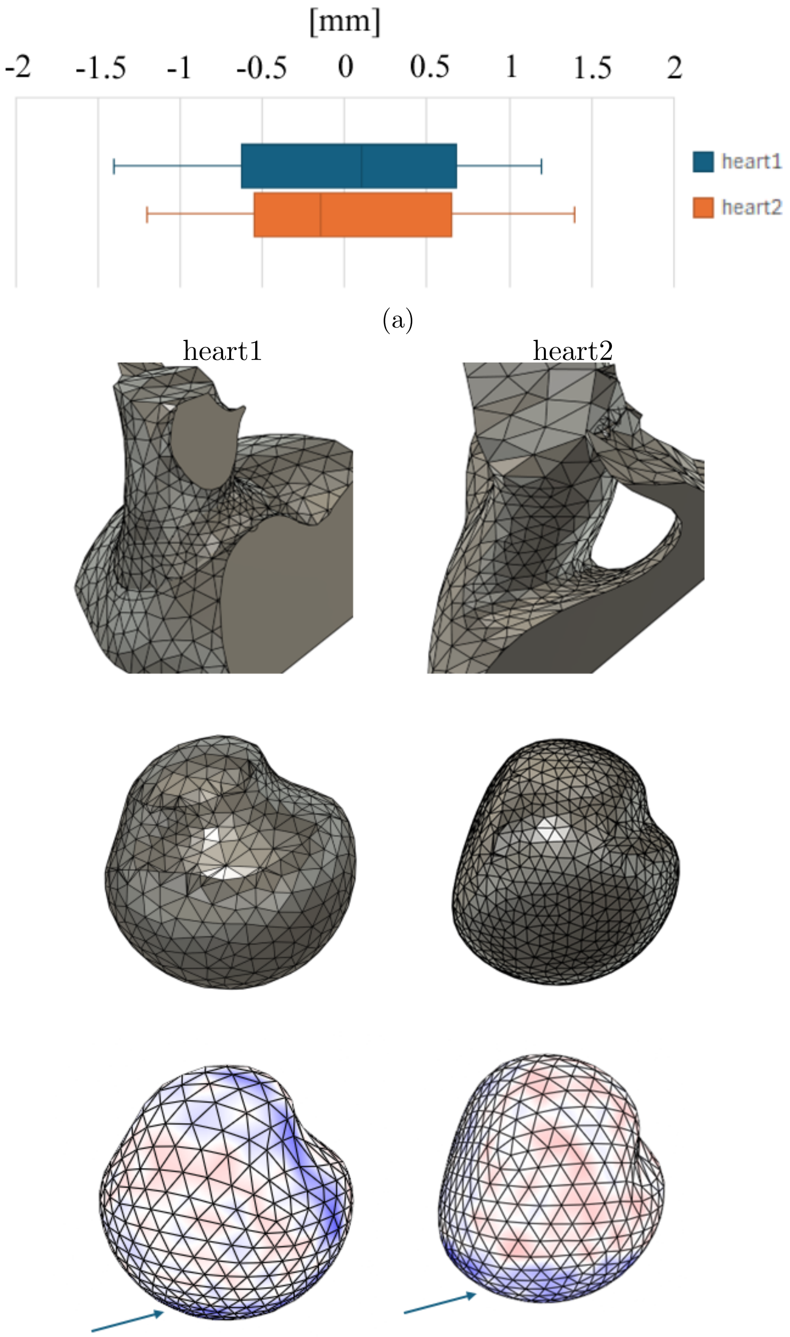


Fig. 4.4: Results of the shape estimation experiment. (a) Euclidean error between the estimated vertex coordinates of the balloon catheter and the ground-truth vertex coordinates obtained when the balloon is inflated along the internal shape. (b) (Top) Target shapes. (Bottom) Euclidean error between estimated and ground-truth vertex positions. White indicates low error; red and blue represent over- and under-estimations, respectively. The arrow points toward the balloon base, outside the endoscope’ s field of view.

Similarly, Fig. 4.5b visualizes, on a per-vertex basis, the distance between the estimated and ground-truth vertices for two shape patterns that emulate contact with the pulmonary veins of the heart.

4.3.3 Discussion

Across all evaluated shapes, the error converged to approximately ±1mm, indicating that the Tactile Balloon Network provides stable 3D shape estimation even for diverse inner-wall geometries. In this evaluation, we tested estimation performance under different deformation patterns, including spherical surfaces, concave regions, bulged regions, and



(a)

(b)

Fig. 4.5: Results of the shape estimation experiment(heart geometry). (a) Euclidean error between the estimated vertex coordinates of the balloon catheter and the ground-truth vertex coordinates obtained when the balloon is inflated along the internal shape. (b) Contact target shape extracted from the pulmonary veins and their ostia of the heart model, together with the error between the estimated balloon shape and the ground truth.

asymmetric shapes. In all cases, accurate reconstruction was achieved, demonstrating the robustness of the proposed approach.

On the other hand, for all cases (Shapes 1–4), Fig. 4.4b shows a localized increase in estimation error on the balloon base side indicated by the arrow, which lies outside the camera’ s field of view. Since this region cannot be directly observed by the endoscopic camera and marker information is missing, the network must interpolate the shape based on surrounding global deformation cues. As a result, estimation accuracy tends to be lower than in regions within the field of view.

Similarly, as shown in Fig. 4.4b, for contact scenarios with the heart model, the overall error also converges to approximately ±1mm, while the error again increases on the base-side region outside the camera view.

These trends are consistent with the observations in the depth-estimation experiments, indicating that missing visual information inherent to vision-based tactile sensing similarly affects shape estimation performance.

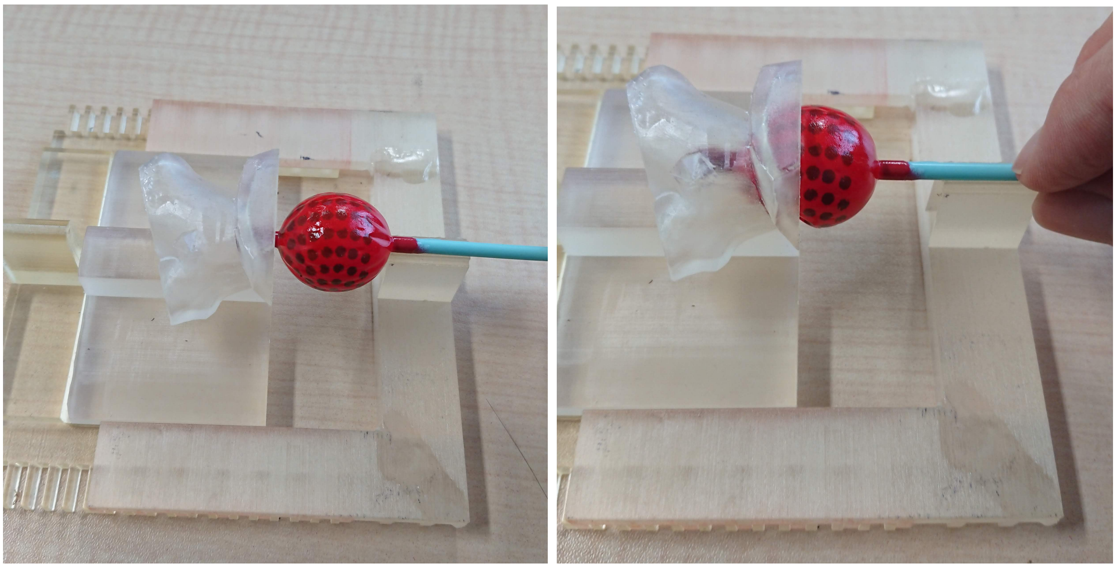

(a) (b)

Fig. 4.6: Experimental setup for contact area estimation. (a) State before balloon insertion: depth = 0 mm, where the balloon tip is aligned with the entrance of the contact object. (b) State after balloon insertion: depth (mm) denotes the insertion depth of the balloon tip.

4.4 Contact Area Estimation Accuracy

In balloon catheter ablation, how accurately and stably the balloon catheter contacts the atrial wall is a critical factor that directly affects the success rate and therapeutic efficacy of the procedure. Inadequate contact may lead to incomplete ablation of the target region, thereby reducing treatment efficacy and increasing the risk of arrhythmia recurrence. Clinical studies have shown that intraoperative monitoring of contact force and contact area significantly contributes to the quality of lesion formation [15, 16]. In practice, the width of ablated tissue created during clinical procedures is typically on the order of 4–12, mm; therefore, a sensing accuracy capable of resolving at least a 4, mm-wide lesion is required [88].

In this section, we evaluate the ability of the proposed vision-based tactile balloon catheter and Tactile Balloon Network to estimate the contact shape and contact area between the balloon and target structures. Specifically, we investigate whether the model can accurately reproduce the three-dimensional deformation patterns that occur when the balloon contacts complex, asymmetric geometries such as the atrial wall and pulmonary veins.

To this end, we fabricated several test models with heart-like geometries using transparent resin and recorded the three-dimensional deformation patterns that arise when the balloon contacts these complex curved surfaces. This experimental framework makes it possible to assess whether the proposed method can reconstruct contact regions not only for simple spherical contacts but also for complex shapes closer to those encountered in clinical practice.

Based on this setup, we evaluate the accuracy of contact shape estimation, the reproducibility of the spatial extent of the contact region, and the quantitative accuracy of contact area estimation, and discuss the effectiveness of the proposed system and its potential for clinical application.

4.4.1 Experimental Design

In the contact shape estimation experiment, we used evaluation objects fabricated by a stereolithography (SLA) 3D printer, as shown in Fig. 4.6. A semi-transparent resin was selected as the printing material so that the contact region between the balloon catheter and the target object could be visually observed from outside. When expanding a balloon catheter inside the heart, only the minimum required amount of fluid is injected to prevent the catheter from bending due to the balloon’ s pressure-induced reaction force and slipping out of the pulmonary vein. Consequently, the balloon does not generate sufficient pressure to displace the myocardium. Therefore, in this study, contact evaluation was performed using rigid resin targets that do not deform.

The target object was fixed to a dedicated jig, and an experimental setup like that shown in Fig. 4.6 was constructed. The tactile balloon catheter was then inserted into the internal cavity of the target object. The posture of the balloon catheter was stabilized by inserting a metallic guidewire—of the type commonly used in clinical ablation procedures —into the inner tube, thereby preventing unwanted rotation or lateral displacement of the catheter inside the target object. Once the catheter tip reached the prescribed insertion depth, fluid was injected into the balloon to inflate it and bring it into contact with the inner wall of the target object. The inflation volume was adjusted such that wrinkles did not appear on the non-contact side, ensuring that the entire balloon surface smoothly followed the inner wall. To minimize the influence of friction-induced deformation (e. g. , pulling or dragging) between the balloon surface and the inner wall and to evaluate the

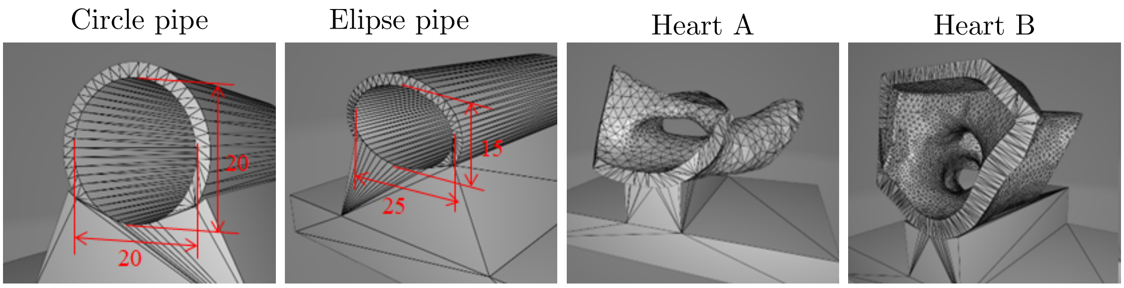


Fig. 4.7: Contact targets for contact area estimation. 3D shapes used for contact estimation evaluation.

pure contact shape, a thin layer of silicone oil was applied to the inner surface of the target object.

The shapes of the target objects used for evaluation are shown in Fig. 4.7. Four types were employed in the experiment:

- (a) Circular pipe: a simple cylindrical shape used for baseline evaluation.
- (b) Elliptical pipe: a pipe with an anisotropic cross-section used to assess the effect of anisotropic deformation direction and magnitude.
- (c) Heart model A: a 3D model of a region around the pulmonary veins, extracted from a mesh obtained from the NIH heart library [87], used to evaluate estimation under complex geometries.
- (d) Heart model B: another 3D model of a different pulmonary vein region from the NIH heart library [87], also used to evaluate estimation under complex geometries.

Heart models A and B reflect characteristic geometrical features of the pulmonary vein ostia and are used as evaluation targets to simulate clinically relevant contact scenarios in balloon ablation.

The contact area estimation experiment was conducted under two insertion depth conditions: 10 mm and 15 mm. For each condition, the tactile balloon catheter was inserted into the target object, inflated, and held in a stable contact state during measurement. This protocol allows comparison of the contact region extent and estimation accuracy under different degrees of insertion and for complex geometries.

4.4.2 Estimation of Contact Area

To obtain data for comparison, we first extracted the regions where the balloon catheter actually contacts the target object. When the balloon surface comes into contact with the target surface, the local reflection conditions change due to compression, and this appears in the images as regions where brightness and color differ from the surrounding areas, as illustrated in Fig. 4.8. In this experiment, we exploit these changes in reflectance to visually extract the contact region.

After the balloon is inflated and brought into contact, the target object is photographed from multiple viewpoints. The resulting image set is then processed by photogrammetry to reconstruct the three-dimensional geometry. Because the contact region on the underside of the object cannot be captured due to occlusion by the supporting base, this portion is excluded from the ground-truth data. Among the reconstructed three-dimensional model, the region exhibiting brightness differences due to changed reflection conditions is designated as the ground-truth contact area for the experiment.

For the Tactile Balloon Network–based estimation, the three-dimensional balloon mesh predicted by the network is virtually inserted into the design model of the contact target at the same insertion depth as in the physical experiment. The minimum distance between vertices on the balloon mesh and the target surface mesh is then computed. Regions where this distance is less than or equal to 0. 5 mm are regarded as contact regions. This threshold is chosen based on the shape estimation accuracy obtained in the previous experiments.

The extracted contact regions can be visualized as three-dimensional meshes, as shown in Fig. 4.8a. Next, cylinder projection is applied with respect to the balloon insertion axis to unfold the three-dimensional contact surface onto a two-dimensional plane. This procedure maps the contact region distributed on a curved surface to a planar representation, which facilitates visual analysis and comparison of contact area and spatial distribution.

Figure 4.9 shows the two-dimensional unfolded maps of the true and estimated contact regions. The blue regions represent the ground-truth contact regions obtained by photogrammetry, while the red regions represent the contact regions extracted from the estimated balloon shapes reconstructed by the Tactile Balloon Network. By overlaying these regions, we can directly compare the consistency of contact location, the extent of the contact region, and the local shape of the contact distribution, thereby evaluating the accuracy and reproducibility of contact area estimation achieved by the proposed method.

4.4.3 Experimental Results and Discussion

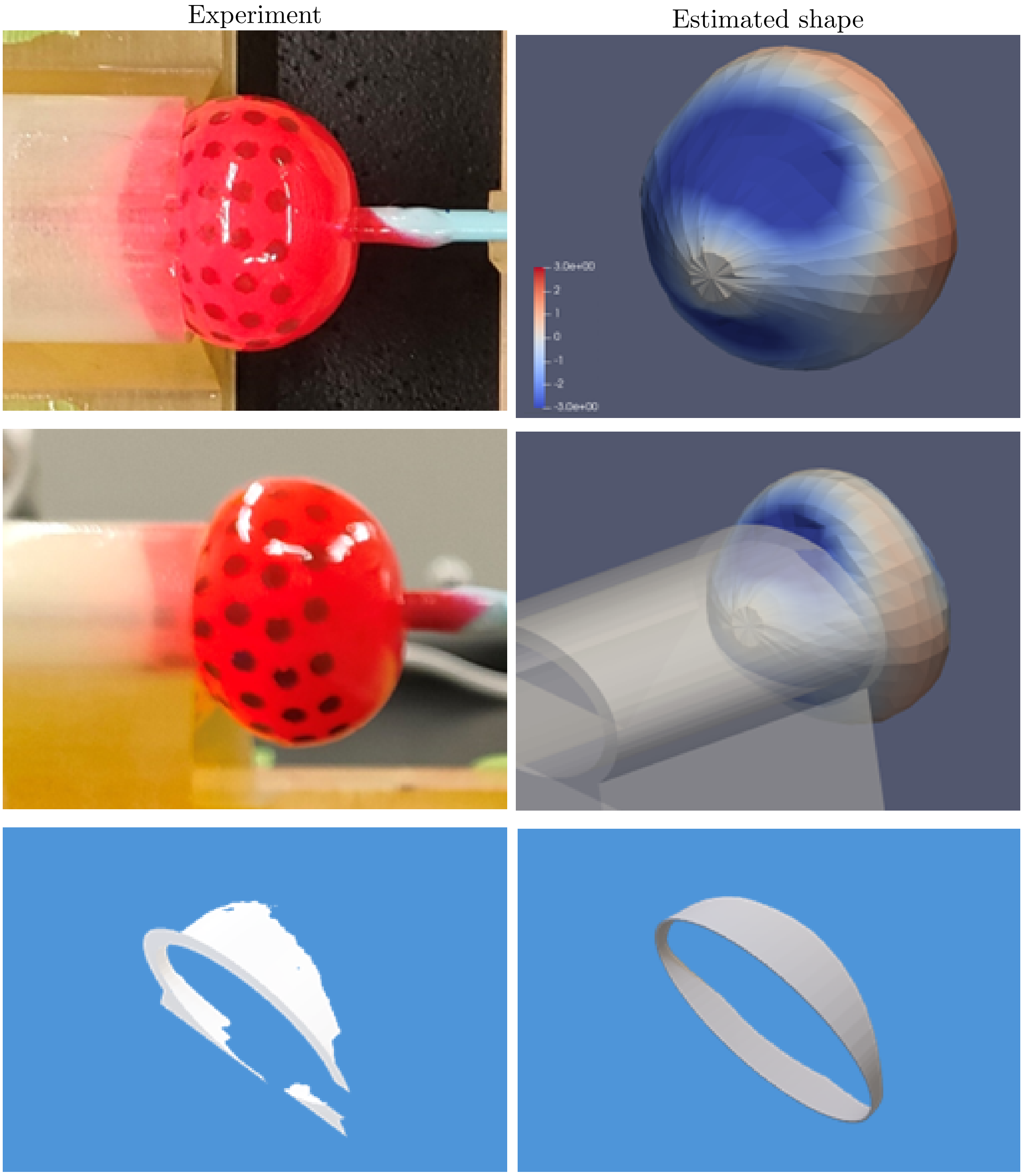
As shown in Fig. 4.9, the estimated contact regions generally agree well with the actual contact areas regardless of target shape or insertion depth. In particular, for the circular and elliptical pipes (Fig. 4.7(a)(b)), the discrepancy in both the shape and extent of the contact region is kept within 1 mm. This indicates that the proposed tactile estimation model exhibits high generalization performance for relatively simple geometries.

On the other hand, for both the circular and elliptical pipes, there are localized areas where the estimated contact area (red) and the measured contact area (blue) exceeds 1 mm, and the estimation accuracy degrades continuously along the balloon' s axial direction. These areas correspond to regions that are occluded by the inner tube and where surface markers cannot be observed within the endoscopic field of view. As confirmed in the depth and shape estimation experiments, such view-deficient regions suffer from degraded estimation accuracy, and this effect manifests as reduced contact area estimation accuracy.

In the cases of Heart A and Heart B in Fig. 4.9, the model is able to follow complex anatomical geometries, such as the highly curved surfaces and branching structures characteristic of the pulmonary veins, and reconstruct the corresponding contact regions. In particular, at an insertion depth of 15 mm, the model identifies contact regions around the branched pulmonary veins. However, for these complex heart-contact scenarios, the estimation error is larger than that for simpler shapes, with a typical discrepancy of approximately 1–2 mm between the estimated (red) and measured (blue) contact areas.

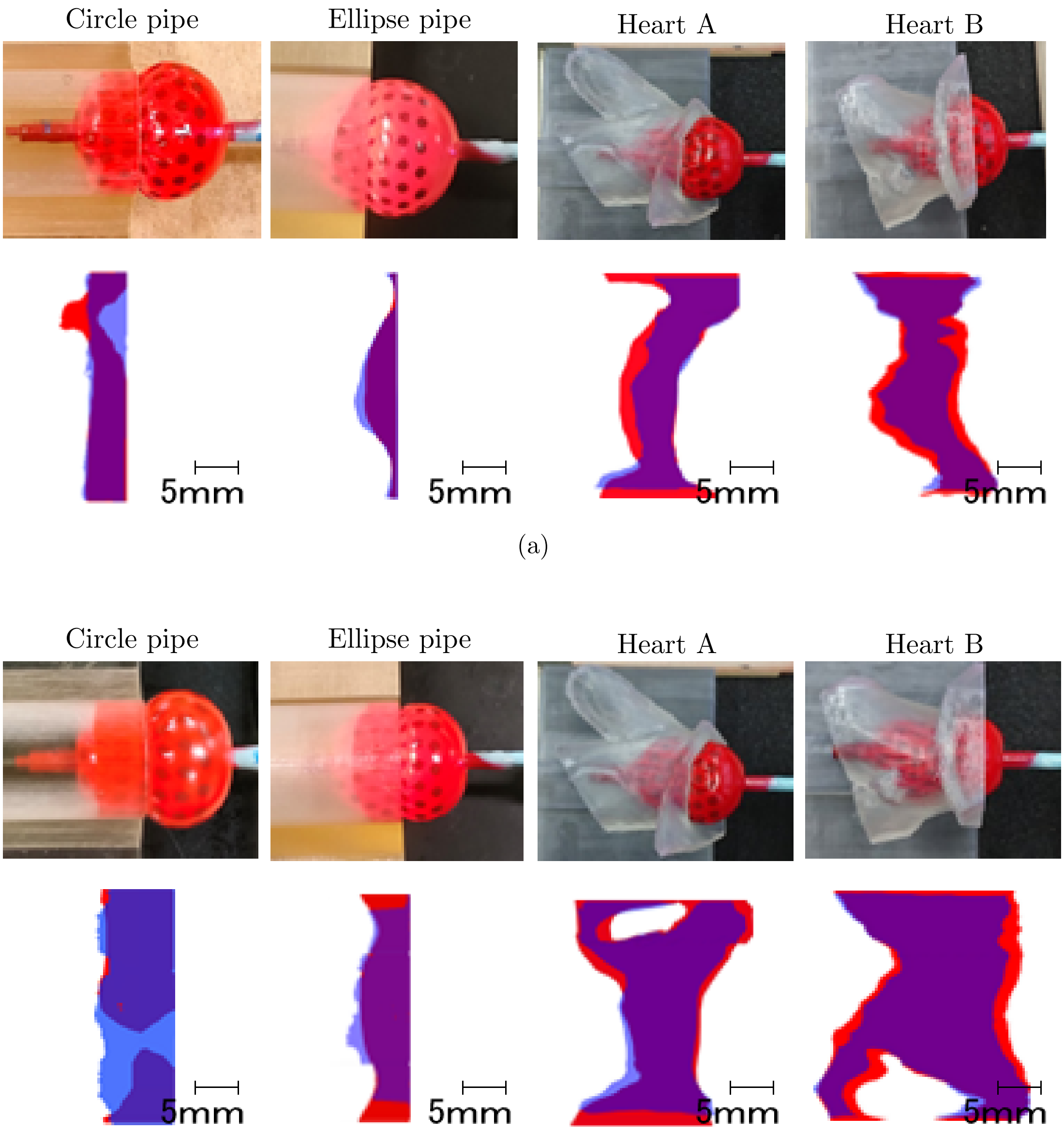
One plausible cause is balloon bending induced by unbalanced contact forces applied from different directions during interaction with the complex heart geometry. Such bending could lead to misalignment between the coordinate frame used for depth-based registration and the actual pose of the balloon relative to the 3D model of the contact object.

Overall, these results confirm that the proposed Tactile Balloon Network can estimate contact regions for both simple geometries and complex heart shapes. At the same time, consistent with the depth- and shape-estimation results, missing visual information—such as marker occlusion and out-of-FOV regions—directly translates into estimation error. Relative to the lower bound of the clinically relevant ablation width (approximately 4 mm), the accuracy obtained for complex geometries corresponds to an error on the order



(a) (b)

Fig. 4.8: Estimation method for the contact region. (Left) Estimation of the contact region based on photogrammetry. The contact region is photographed from top and side views and reconstructed into 3D by photogrammetry. The region whose appearance changes due to contact is extracted as the ground-truth contact area. (Right) Estimation of contact region based on the tactile model: the reconstructed balloon shape is inserted into the CAD model of the contact target under the same conditions as in the experiment, and the region where the distance between the meshes falls below a predefined threshold is extracted as the estimated contact area.



(a)

(b)

Fig. 4.9: Results of contact area estimation. Blue regions represent ground-truth contact areas obtained from 3D shapes reconstructed by photogrammetry, and red regions represent estimated contact areas extracted from the reconstructed balloon shapes. (a) Contact areas at 10 mm insertion depth. (b) Contact areas at 15 mm insertion depth.

of 50 %.

Chapter5

General Discussion

5.1 Discussion

In the field of medical devices, where available space is severely limited, tactile sensing has traditionally relied primarily on electrical sensing principles. However, electrical tactile sensors face fundamental constraints when it comes to maintaining flexibility and realizing wide-area, surface-level tactile sensing. In this study, we addressed these challenges by proposing a balloon-catheter-type tactile sensor that combines optical markers drawn on the balloon surface with image-based machine learning inference, thereby enabling the acquisition of wide-area tactile information without any wiring or electrodes on the sensing surface.

Furthermore, the development of ultra-compact vision-based tactile sensors such as the proposed balloon catheter–type sensor is difficult to achieve using conventional tactile sensor development methods, especially under strict geometric and mechanical constraints. To simplify the development process and improve its flexibility, we constructed a tactile sensing system in which the tactile estimation model is trained using a large-scale dataset generated entirely in a simulation environment. As a concrete implementation, we designed a vision-based tactile sensing method (TacBalloon) that combines an endoscopic camera and optical markers on a balloon catheter commonly used in cardiac ablation procedures. We then implemented a mechanism to estimate balloon shape and contact state using the machine learning model Tactile Balloon Network. Through a series of depth estimation, shape estimation, and contact area estimation experiments, we confirmed that TacBalloon can function effectively as a tactile sensing modality for medical applications.

To mitigate the Sim2Real gap, we applied domain randomization of physical parameters in the simulation and performed data augmentation with artificial noise during training. These strategies enabled the tactile estimation model, trained solely on simulation data, to operate robustly in real environments. As a result, we achieved a depth estimation error of approximately ±1 mm in visible regions within the field of view, and obtained high reproducibility for both global balloon shape estimation and contact area estimation.

In addition, the proposed method is not limited to balloon catheters; it is applicable to a wide range of sensor structures such as multi-degree-of-freedom robotic hands and flexible finger-type sensors, where endoscopic camera fixation is inherently difficult. The ability to freely modify the sensor geometry is also a major advantage, making it straightforward to design marker patterns and optimize sensing performance according to specific applications.

5.2 Limitations of Sensing Capability

While TacBalloon demonstrated high accuracy in depth, shape, and contact area estimation within the visible field of view, several limitations in performance were identified. The most prominent factor is the decrease in estimation accuracy in regions outside the endoscope field of view and in areas occluded by the inner shaft. In out-of-view regions, markers cannot be imaged and thus input information is missing. Consequently, the model is forced to infer deformation by globally interpolating from surrounding visible regions, making it difficult to reconstruct localized contact or deep deformations. Similarly, in occluded regions, markers are partially lost, leading to a lack of the features required by the model and, ultimately, to increased estimation error.

Moreover, in real environments, large deformation of the balloon surface can cause markers to become severely distorted or overlap, which destabilizes the extraction of marker positions. In addition, optical noise such as scattering from underwater environments and internal reflections within transparent resin—which are difficult to reproduce perfectly in simulation—also affects visual information and tends to exacerbate estimation errors.

In this study, we addressed these issues by introducing data augmentation that included various types of noise and partial occlusions. However, it is inherently difficult to completely cover the diversity of real-world optical conditions. In the contact area estimation, we adopted a strict criterion that defined contact as the region where the minimum distance between meshes is less than or equal to 0. 5 mm. As a result, even small shape estimation errors or slight axial misalignment of the balloon sometimes manifested as over- or under-estimation of the contact region.

These observations indicate that the estimation performance of TacBalloon strongly depends on marker visibility and the stability of the imaging environment. To further improve robustness against these limitations, several possible approaches can be considered:

- **Diversification of input modalities:** In addition to endoscopic images, incorporating sensor values related to the internal fluid pressure and inflation volume of the balloon as additional inputs would enable multimodal estimation of the deformation state that integrates both geometric and mechanical information. Such input fusion is expected to reduce the dependence of estimation accuracy on image quality alone, thereby mitigating performance degradation in out-of-view or occluded regions.

- **Stabilization of endoscopic camera pose:** In the present structure, the camera inserted at the balloon center can move and tilt, and such visual fluctuations can lead to increased estimation errors. Mechanically fixing the camera to the balloon base and reducing the deviation between the optical axis and the catheter axis would improve imaging stability and potentially enhance estimation accuracy. However, irreversible fixation using adhesives is economically unfavorable because it couples the lifetime of the endoscopic camera (a reusable component) to that of the balloon catheter (a disposable device). On the other hand, implementing a mechanically reversible fixation mechanism within the current catheter diameter is technically challenging.

- **Optimization of marker patterns:** By optimizing marker density, shape, color contrast, and spatial distribution according to typical contact patterns and deformation modes, it may be possible to design encoding patterns that are simultaneously sensitive to both local and global deformation.

- **Higher-fidelity simulation:** Incorporating more accurate models of soft-body mechanics, such as hyperelasticity, into the soft-body simulation used for dataset construction, and generating datasets under diverse elastic properties, friction conditions, and deformation behaviors, could further reduce the Sim2Real gap and strengthen generalization performance in real environments.

Furthermore, in this study, contact was indirectly inferred from shape deformation rather than directly estimating the contact force itself. Methods that estimate contact force based on changes in balloon internal pressure [89] or by inverse analysis from deformation shapes [90, 91] have been proposed. Integrating such approaches with TacBalloon could extend the device into a tactile sensor capable of simultaneously sensing both shape and force.

Table 5.1: Comparison of vision-based tactile endoscopes

Item	Gelsight type colonoscope Kim *et al.* [77]	TacBalloon
Minimum diameter [mm]	25	5
Maximum diameter [mm]	90	30
Sensing region	40 mm × 25 mm (torus outer surface)	Entire balloon (approximately spherical)
Spatial resolution	$\approx \pm 53\,\mu$m	$\approx \pm 1$ mm (within FoV)
Structural design	Rigid structure with integrated optical module	Fully flexible balloon integrated structure
Scalability	Limited	Shape and size scalable

5.3 Comparison with Vision-Based Tactile Endoscopes

In this section, we compare the proposed TacBalloon with a representative existing vision-based tactile endoscope system, namely the GelSight-sensor-equipped colonoscopic system proposed by Kim et al. [77], and discuss their respective characteristics and advantages.

We first focus on device geometry. In the GelSight-based tactile colonoscope, the structural constraints of the optical module result in a relatively large minimum diameter of approximately 25 mm. Upon inflation, the balloon can expand up to about 90 mm in outer diameter. This allows for a wide contact area, but limits the anatomical regions into which the device can be safely inserted. In contrast, TacBalloon utilizes an existing ablation balloon catheter structure directly as a tactile sensor, maintaining a compact profile with a minimum diameter of approximately 5 mm and a maximum diameter of approximately 30 mm. This compactness enables insertion into narrow and complex clinical environments such as pulmonary veins and the atrial chamber, giving TacBalloon a broader range of potential applications than the GelSight-based system. Moreover, the TacBalloon design is inherently scalable, allowing the balloon size to be modified to suit various applications and sensor configurations.

Next, we compare the sensing regions. In the GelSight-based colonoscope, the effective sensing region is limited to a portion of the outer surface of a toroidal balloon, with an effective area of approximately 40 mm × 25 mm [77]. In contrast, TacBalloon can be patterned with optical markers over the entire balloon surface, enabling tactile sensing over the full circumference without geometric restriction. This is especially advantageous in procedures such as cardiac ablation, where circumferential and wide contact with the tissue is required.

From the viewpoint of measurement accuracy, GelSight has a clear advantage in high-resolution measurement of fine surface deformations, achieving an axial resolution of approximately $\pm 53\,\mu$m within the sensing region [77]. This high spatial resolution is particularly beneficial for applications such as surface texture analysis and detection of subtle topographical features, making it well-suited for diagnostic use. On the other hand, TacBalloon achieves a depth estimation accuracy of about ± 1 mm in unoccluded regions within the field of view, which is lower in resolution compared to GelSight.

In summary, the GelSight-based endoscope excels at acquiring high-resolution, localized tactile information, but its device size and sensing region are constrained. In contrast, TacBalloon offers superior compactness, flexibility, and wide-area sensing capability, making it particularly suitable for use in narrow and complex anatomical environments such as those encountered in cardiac ablation. Thus, the two approaches have complementary strengths, and the appropriate choice depends on the measurement target and clinical application.

Chapter6

Conclusion

6.1 Summary

In this thesis, we proposed a novel vision-based tactile sensor, TacBalloon, specifically designed for minimally invasive surgery, and demonstrated its effectiveness through its design, learning framework, and experimental evaluation. TacBalloon integrates optical markers and an endoscopic camera into a balloon catheter widely used in cardiac ablation procedures. By acquiring three-dimensional deformation of the balloon surface as visual information, it enables tactile sensing that simultaneously satisfies flexibility, wide-area coverage, and compact form factor—requirements that are difficult to achieve with conventional electrically based tactile sensors.

The proposed method has the advantage that a tactile estimation model can be constructed without prototyping and measuring a physical tactile sensor. This is achieved by using large-scale datasets of balloon deformation behaviors generated by physics-based simulation as training data for supervised learning. The proposed Tactile Balloon Network estimated the three-dimensional shape of the balloon from changes in the marker patterns in endoscopic images, and achieved a high depth estimation accuracy of approximately ± 1 mm within the field of view in depth estimation experiments. In shape estimation experiments, it exhibited stable reconstruction performance for a wide range of target geometries, from spherical shapes to complex cardiac geometries, and in contact area estimation experiments, it showed good agreement with experimentally measured data. These results indicate that TacBalloon has the potential to practically reconstruct tactile information even in complex environments that approximate actual clinical conditions.

Moreover, the proposed approach is not limited to balloon catheters. By combining visual information with flexible structures, it can be extended to a wide variety of medical devices and soft robotic systems, significantly increasing the design freedom of tactile sensing. Since it does not require additional electrical wiring on the sensing surface, it is particularly advantageous in medical applications where high safety and durability are required.

6.2 Future Work

Although the proposed TacBalloon has demonstrated its effectiveness as a new tactile sensing system, further technical developments are required for clinical application. In particular, the strong dependence of the method on marker observations within the field of view is a fundamental factor underlying the instability of tactile estimation in out-of-view or occluded regions, and overcoming this limitation is a key challenge for future work. We expect that stabilizing the camera pose, optimally designing the marker patterns, and introducing more advanced model architectures that incorporate temporal information and self-supervised learning will help compensate for missing visual information and improve robustness.

Furthermore, to further advance TacBalloon, multimodal tactile estimation that integrates information sources other than vision will be essential. For example, by incorporating information related to changes in the internal pressure of the balloon or the membrane tension, it may become possible to estimate not only the contact geometry but also the contact force, thereby realizing more comprehensive tactile sensing. In addition, since the training data for TacBalloon rely on simulation, improving the accuracy of soft-body simulation is also an important issue.

TacBalloon was implemented on a balloon catheter designed with cardiac ablation as a specific application in mind. However, its structural design and processing framework are inherently scalable, and thus hold potential for application to other medical devices and soft robotic systems. Possible future applications include integration into colonoscopic devices, soft robotic hands, and large-area robotic skins, among others.

Acknowledge

I would like to express my sincere gratitude to Associate Professor Ho Anh Van of the School of Materials Science,Graduate School of Advanced Science and Technology,Japan Advanced Institute of Science and Technology,for his warm guidance and numerous valuable insights throughout the course of this research. His advice has been invaluable in shaping both the direction of the study and my academic perspective.

I also wish to extend my heartfelt thanks to all collaborators at Toray Industries,Inc. ,including the members of the Cardiac Ablation Catheter Development Office,who have greatly supported this work as part of a joint research project. Their provision of experimental environments,technical expertise,and continual cooperation were essential to the progress of this study.

This research was supported by the Japan Science and Technology Agency’ s Support for Pioneering Research Initiated by the Next Generation (SPRING) program (JPMJSP2102). I gratefully acknowledge this support.

I am additionally indebted to the members of my laboratory for their continuous encouragement and for the insightful discussions that enriched my research activities on a daily basis.

Finally,I would like to express my deepest appreciation to my parents,whose understanding and unwavering support made it possible for me to pursue doctoral studies.

Reference

[1] “平成21年度ロボット産業の新規市場創出に向けた国内外技術動向及び市場分析に係る情報収集成果報告書”, Tech. Rep. 20100000001144, New Energy and Industrial Technology Development Organization (NEDO) and Mitsubishi Research Institute Inc., May 2020.

[2] ABI Research, “Commercial and industrial robotics addressable market, segmentation, and technologies market data overview: 3q 2025”, July 2025.

[3] Future Market Insights, “Industrial robotics market outlook 2025 to 2035”, February 2025, `https://web.archive.org/web/20251202185741/https://www.futuremarketinsights.com/reports/industrial-robotics-market`.

[4] Mordor Intelligence, “Industrial robotics market size share analysis - growth trends forecasts (2025 - 2030)”, June 2025, `https://web.archive.org/web/20250621232713/https://www.mordorintelligence.com/industry-reports/industrial-robotics-market`.

[5] Future Market Insights, “Service robotics market size and share forecast outlook 2025 to 2035”, August 2025, `https://web.archive.org/web/20251202185821/https://www.futuremarketinsights.com/reports/service-robotics-market`.

[6] 岸本喜久雄, 植木健司, 伊藤智, 桜井茂行, 吉野順也, 仙洞田充, 林 秀樹, 高木宗谷, 山口佳樹, 麻生英樹, 尾形哲也, and 加藤浩瑞, “人工知能×ロボット分野の技術戦略策定に向けて”, Tech. Rep. TSCForesight 120, New Energy and Industrial Technology Development Organization Technology Strategy Center (NEDO TSC), June 2024, `https://www.nedo.go.jp/content/100978754.pdf`.

[7] Science of Soft Robots: Interdiscplinary integration of mechatronics, material science, and bio-computing KAKENHI Grant-in-Aid for Scientific Research on Innovative

Areas (2018-2022), “Science of soft robots/outline of the area”, `https://softrobot.jp/en/outline/`.

[8] Chunfeng Wang, Lin Dong, Dengfeng Peng, and Caofeng Pan, “Tactile sensors for advanced intelligent systems”, *Advanced Intelligent Systems*, vol. 1, no. 8, pp. 1900090, August 2019, `https://doi.org/10.1002/aisy.201900090`.

[9] “Tactile sensors: A review”, *Measurement*, vol. 238, pp. 115332, October 2024, `https://doi.org/10.1016/j.measurement.2024.115332`.

[10] Peter Roberts, Mason Zadan, and Carmel Majidi, “Soft tactile sensing skins for robotics”, *Current Robotics Reports*, vol. 2, pp. 343–354, July 2021, `https://doi.org/10.1007/s43154-021-00065-2`.

[11] Micah K. Johnson and Edward H. Adelson, “Retrographic sensing for the measurement of surface texture and shape”, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 2009, pp. 1070–1077, IEEE, `https://doi.org/10.1109/CVPR.2009.5206534`.

[12] Benjamin Ward-Cherrier, Nicholas Pestell, Luke Cramphorn, Benjamin Winstone, Maria Giannaccini, Jonathan Rossiter, and Nathan Lepora, “The tactip family: Soft optical tactile sensors with 3d-printed biomimetic morphologies”, *Soft Robotics*, vol. 5, no. 2, pp. 216–227, January 2018, `https://doi.org/10.1089/soro.2017.0052`.

[13] Nathan F. Lepora, “Soft biomimetic optical tactile sensing with the tactip: A review”, *IEEE Sensors Journal*, vol. 21, no. 19, pp. 21131–21143, Octber 2021, `https://doi.org/10.1109/JSEN.2021.3100645`.

[14] Lac Van Duong and Van Anh Ho, “Large-scale vision-based tactile sensing for robot links: Design, modeling, and evaluation”, *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 390–403, April 2021, `https://doi.org/10.1109/TRO.2020.3031251`.

[15] Jackson J. Liang and Pasquale Santangeli, “Contact force sensing during atrial fibrillation ablation: clinical experience and effects on outcomes”, *Expert Review of Cardiovascular Therapy*, vol. 14, no. 6, pp. 749–759, March 2016, `https://doi.org/10.1586/14779072.2016.1168695`.

[16] Vivek Y. Reddy, Dipen Shah, Josef Kautzner, Boris Schmidt, Nadir Saoudi, Claudia Herrera, Pierre Jaïs, Gerhard Hindricks, Petr Peichl, Aude Yulzari, Hendrik Lambert, Petr Neuzil, Andrea Natale, and Karl-Heinz Kuck, "The relationship between contact force and clinical outcome during radiofrequency catheter ablation of atrial fibrillation in the toccata study", *Heart rhythm : the official journal of the Heart Rhythm Society*, vol. 9, no. 11, pp. 1789–1795, Novenber 2012, `https://doi.org/10.1016/j.hrthm.2012.07.016`.

[17] V. S. N. Sitaramgupta V., Deepak Padmanabhan, Prasanna Simha Mohan Rao, and Hardik J. Pandya, "Force sensing technologies for catheter ablation procedures", *Mechatronics*, vol. 64, no. 102295, December 2019, `https://doi.org/10.1016/j.mechatronics.2019.102295`.

[18] So-Hyun Cho, Su-Min Lee, Na-Young Lee, Byoung Chul Ko, Hojeong Kim, Dae-Jin Jang, and Jong-Ha Lee, "High-resolution tactile-sensation diagnostic imaging system for thyroid cancer", *Sensors*, vol. 23, no. 7, pp. 3451, March 2023, `https://doi.org/10.3390/s23073451`.

[19] Siddhartha Kapuria, Jeff Bonyun, Yash Kulkarni, Naruhiko Ikoma, Sandeep Chinchali, and Farshid Alambeigi, "Robot-enabled machine learning-based diagnosis of gastric cancer polyps using partial surface tactile imaging", in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, United Arab Emirates, 2024, pp. 2360–2365, IEEE, `https://doi.org/10.1109/IROS58592.2024.10802585`.

[20] Chenxia Wu, Luoxia Hu, Youjin Kong, Bowen Zhao, Wei Mao, and Xinbin Zhou, "Bayesian network meta-analysis comparing hot balloon, laser balloon and cryoballoon ablation as initial therapies for atrial fibrillation", *Frontiers in Cardiovascular Medicine*, vol. 10, no. 1184467, July 2023, `https://doi.org/10.3389/fcvm.2023.1184467`.

[21] Shiro Nakahara, Yuichi Hori, Reiko Fukuda, Hirotsugu Sato, Hideyuki Aoki, Yuki Kondo, Yuta Kimura, Yuji Itabashi, Tetsuya Ishikawa, Sayuki Kobayashi, and Isao Taguchi, "Current balloon devices for ablation of atrial fibrillation", *Rev Cardiovasc Med*, vol. 25, no. 1, pp. 0–12, January 2024, `https://doi.org/10.31083/j.rcm2501034`.

[22] Nebojša Mujović, Milan Marinković, Radoslaw Lenarczyk, Roland Tilz, and Tatjana S. Potpara, “Catheter ablation of atrial fibrillation: An overview for clinicians”, *Advances in Therapy*, vol. 34, no. 8, pp. 1897–1917, August 2017, `https://doi.org/10.1007/s12325-017-0590-z`.

[23] Florian Straube, Uwe Dorwarth, Janis Pongratz, Benedikt Brück, Michael Wankerl, Stefan Hartl, and Ellen Hoffmann, “The fourth cryoballoon generation with a shorter tip to facilitate real-time pulmonary vein potential recording: Feasibility and safety results”, *Journal of Cardiovascular Electrophysiology*, vol. 30, no. 6, pp. 918–925, June 2019, `https://doi.org/10.1111/jce.13927`.

[24] Shutaro Satake, Kazushi Tanaka, Shigeru Saito, Shinji Tanaka, Hiroshi Sohara, Yoshitaka Hiroe, Yusuke Miyashita, Saeko Takahashi, Masato Murakami, and Yoshio Watanabe, “Usefulness of a new radiofrequency thermal balloon catheter for pulmonary vein isolation:. a new device for treatment of atrial fibrillation”, *Journal of cardiovascular electrophysiology*, vol. 14, no. 6, pp. 609–615, June 2003, `https://doi.org/10.1046/j.1540-8167.2003.02577.x`.

[25] Kazushi Tanaka, Shutaro Satake, Shigeru Saito, Saeko Takahashi, Yoshitaka Hiroe, Yusuke Miyashita, Shinji Tanaka, Michio Tanaka, and Yoshio Watanabe, “A new radiofrequency thermal balloon catheter for pulmonary vein isolation”, *Journal of the American College of Cardiology*, vol. 38, no. 7, December 2001, `https://doi.org/10.1016/S0735-1097(01)01666-7`.

[26] Boris Schmidt, Jan Petrů, Kr Chun, Lucie Sediva, Stefano Bordignon, Shaojie Chen, and Petr Neuzil, “Pivotal study of a novel motor driven endoscopic ablation system”, *Circulation: Arrhythmia and Electrophysiology*, vol. 14, no. 3, March 2021, `https://doi.org/10.1161/CIRCEP.120.009544`.

[27] Michifumi Tokuda, Seigo Yamashita, Satoko Shiomi, Ryutaro Sakurai, Hidenori Sato, Hirotsuna Oseto, Masaaki Yokoyama, Kenichi Tokutake, Mika Kato, Ryohsuke Narui, Shin ichi Tanigawa, Michihiro Yoshimura, and Teiichi Yamane, “Pulmonary vein stenosis after catheter ablation of atrial fibrillation using a cryoballoon, hot balloon, or laser balloon”, *Circulation Journal*, vol. 87, no. 12, pp. 1711–1719, November 2023, `https://doi.org/10.1253/circj.CJ-23-0048`.

[28] Medtronic, “Diamondtemp ablation catheter”, `https://web.archive.org/web/20250801124605/https://www.medtronic.com/en-us/healthcare-professionals/products/surgical-energy/ablation/radiofrequency-ablation/catheters/diamondtemp-ablation-catheter.html`.

[29] Medtronic, “Arctic front advance pro cardiac cryoablation catheter”, `https://web.archive.org/web/20251202185201/https://www.medtronic.com/en-us/healthcare-professionals/products/surgical-energy/ablation/cryoablation/catheters/arctic-front-advance-pro-cardiac-cryoablation-catheter.html`.

[30] Toray industries,inc., “Satake・hotballon”, `https://web.archive.org/web/20251202185122/https://www.atrial-fibrillation.toray/medical/`.

[31] CardioFocus, “The heartlightx3 system”, `https://web.archive.org/web/20251202185241/https://cardiofocus.com/technology/-heartlight-x3`.

[32] B.F. Gonçalves, J. Oliveira, P. Costa, V. Correia, P. Martins, G. Botelho, and S. Lanceros-Mendez, “Development of water-based printable piezoresistive sensors for large strain applications”, *Composites Part B: Engineering*, vol. 112, pp. 344–352, March 2017, `https://doi.org/10.1016/j.compositesb.2016.12.047`.

[33] Nanshu Lu, Chi Lu, Shixuan Yang, and John Rogers, “Highly sensitive skin-mountable strain gauges based entirely on elastomers”, *Advanced Functional Materials*, vol. 22, no. 19, pp. 4044–4050, June 2012, `https://doi.org/10.1002/adfm.201200498`.

[34] Tsuyoshi Sekitani, Hiroyoshi Nakajima, Hiroki Maeda, Takanori Fukushima, Takuzo Aida, Kenji Hata, and Takao Someya, “Stretchable active-matrix organic light-emitting diode display using printable elastic conductors”, *Nature materials*, vol. 8, pp. 494–499, May 2009, `https://doi.org/10.1038/nmat2459`.

[35] Hu Liu, Jiachen Gao, Wenju Huang, Kun Dai, Guoqiang Zheng, Chuntai Liu, Changyu Shen, Jiang Guo, Xingru Yan, and Zhanhu Guo, “Electrically conductive strain sensing polyurethane nanocomposites with synergistic carbon nanotubes and graphene bifillers”, *Nanoscale*, vol. 8, pp. 12977–12989, May 2016, `https://doi.org/10.1039/C6NR02216B`.

[36] Samuel Rosset, Muhamed Niklaus, Philippe Dubois, and Herbert R. Shea, “Metal ion implantation for the fabrication of stretchable electrodes on elastomers”, *Advanced Functional Materials*, vol. 19, no. 3, pp. 470–478, February 2009, `https://doi.org/10.1002/adfm.200801218`.

[37] Minwoo Park, Jungkyun Im, Minkwan Shin, Yuho Min, Jaeyoon Park, Heesook Cho, Soojin Park, Mun-Bo Shim, Sanghun Jeon, Dae-Young Chung, Jihyun Bae, Jongjin Park, Unyong Jeong, and Kinam Kim, “Highly stretchable electric circuits from a composite material of silver nanoparticles and elastomeric fibres”, *Nature nanotechnology*, vol. 7, pp. 803–809, Novenber 2012, `https://doi.org/10.1038/nnano.2012.206`.

[38] Ying-Chih Lai, Bo-Wei Ye, Chun-Fu Lu, Chien-Tung Chen, Meng-Huan Jao, Wei-Fang Su, Wen-Yi Hung, Tai-Yuan Lin, and Yang-Fang Chen, “Extraordinarily sensitive and low-voltage operational cloth-based electronic skin for wearable sensing and multifunctional integration uses: A tactile-induced insulating-to-conducting transition”, *Advanced Functional Materials*, vol. 26, no. 8, pp. 1286–1295, January 2016, `https://doi.org/10.1002/adfm.201503606`.

[39] Yancheng Wang, Jie Jin, Yingtong Lu, and Deqing Mei, “3d printing of liquid metal based tactile sensor for simultaneously sensing of temperature and forces”, *International Journal of Smart and Nano Materials*, vol. 12, no. 3, pp. 269–285, April 2021, `https://doi.org/10.1080/19475411.2021.1948457`.

[40] Chwee-Lin Choong, Mun-Bo Shim, Byoung-Sun Lee, Sanghun Jeon, Dong-Su Ko, Tae-Hyung Kang, Jihyun Bae, Sung Hoon Lee, Kyung-Eun Byun, Jungkyun Im, Yong Jin Jeong, Chan Eon Park, Jong-Jin Park, and U-In Chung, “Highly stretchable resistive pressure sensors using a conductive elastomeric composite on a micropyramid array”, *Advanced Materials*, vol. 26, no. 21, pp. 3451–3458, February 2014, `https://doi.org/10.1002/adma.201305182`.

[41] Changhyun Pang, Gil-Yong Lee, Tae il Kim, Sang Moon Kim, Hong Nam Kim, Sung hoon Ahn, and Kahp Y. Suh, “A flexible and highly sensitive strain-gauge sensor using reversible interlocking of nanofibres”, *Nature materials*, vol. 11, pp. 795–801, July 2012, `https://doi.org/10.1038/nmat3380`.

[42] Weizhuang Gong, Jinhong Lian, and Yinlong Zhu, "Capacitive flexible haptic sensor based on micro-cylindrical structure dielectric layer and its decoupling study", *Measurement*, vol. 223, no. 113785, December 2023, `https://doi.org/10.1016/j.measurement.2023.113785`.

[43] Canan Dagdeviren, Yewang Su, Pauline Joe, Raissa Yona, Yuhao Liu, Yun-Soung Kim, Yongan Huang, Anoop Damadoran, Jing Xia, Lane Martin, Yonggang Huang, and John Rogers, "Conformable amplified lead zirconate titanate sensors with enhanced piezoelectric response for cutaneous pressure monitoring", *Nature communications*, vol. 5, no. 4496, August 2014, `https://doi.org/10.1038/ncomms5496`.

[44] Chunfeng Wang, Rongrong Bao, Kun Zhao, Taiping Zhang, Lin Dong, and Caofeng Pan, "Enhanced emission intensity of vertical aligned flexible zno nanowire/p-polymer hybridized led array by piezo-phototronic effect", *Nano Energy*, vol. 14, pp. 364–371, May 2015, `https://doi.org/10.1016/j.nanoen.2014.11.033`.

[45] Shu-Ju Tsai, Chun-Yeh Lin, Chiang-Lun Wang, Jhih-Wei Chen, Chia-Hao Chen, and Chung-Lin Wu, "Efficient coupling of lateral force in gan nanorod piezoelectric nanogenerators by vertically integrated pyramided si substrate", *Nano Energy*, vol. 37, pp. 260–267, July 2017, `https://doi.org/10.1016/j.nanoen.2017.05.033`.

[46] Guo Tian, Weili Deng, Yuyu Gao, Da Xiong, Cheng Yan, Xuebing He, Tao Yang, Long Jin, Xiang Chu, Haitao Zhang, Wei Yan, and Weiqing Yang, "Rich lamellar crystal baklava-structured pzt/pvdf piezoelectric sensor toward individual table tennis training", *Nano Energy*, vol. 59, pp. 574–581, May 2019, `https://doi.org/10.1016/j.nanoen.2019.03.013`.

[47] Mark Daniel Alea, Ali Safa, Flavio Giacomozzi, Andrea Adami, Inci Rüya Temel, Maria Atalaia Rosa, Leandro Lorenzelli, and Georges Gielen, "A fingertip-mimicking 12×16 200 μm-resolution e-skin taxel readout chip with per-taxel spiking readout and embedded receptive field processing", *IEEE Transactions on Biomedical Circuits and Systems*, vol. 18, no. 6, pp. 1308–1320, December 2024, `https://doi.org/10.1109/TBCAS.2024.3387545`.

[48] H. Shinoda, K. Matsumoto, and S. Ando, "Tactile sensing based on acoustic resonance tensor cell", in *Proceedings of International Solid State Sensors and Actuators*

Conference (Transducers '97), Chicago, IL, USA, 1997, vol. 1, pp. 129–132, IEEE, `https://doi.org/10.1109/SENSOR.1997.613599`.

[49] Manuchehr Soleimani and Tomasz Rymarczyk, "A tactile skin system for touch sensing with ultrasound tomography", *Sensors*, vol. 23, no. 13, pp. 6071, July 2023, `https://doi.org/10.3390/s23136071`.

[50] Arman Goshtasbi, Saravana Prashanth Murali Babu, Rita Moreira, Masoud Jamshidiyan Tehrani, Rebecca Hyldgaard, Maria Nikoline Rasmussen, and Ahmad Rafsanjani, "Ai-infused soft fluidic tactile sensing", in *2024 IEEE 7th International Conference on Soft Robotics (RoboSoft)*, San Diego, CA, USA, 2024, pp. 1095–1100, IEEE, `https://doi.org/10.1109/RoboSoft60065.2024.10522049`.

[51] Shibo Zou, Sergio Picella, Jelle Vries, Vera Kortman, Aimee Sakes, and Johannes Overvelde, "A retrofit sensing strategy for soft fluidic robots", *Nature Communications*, vol. 15, no. 539, January 2024, `https://doi.org/10.1038/s41467-023-44517-z`.

[52] Joo Chuan Yeo, Jiahao Yu, Zhao Koh, Zhiping Wang, and C.T. Lim, "Wearable tactile sensor based on flexible microfluidics", *Lab Chip*, vol. 17, pp. 3244–3250, July 2016, `https://doi.org/10.1039/C6LC00579A`.

[53] Suhyeon Kim, Hyeonsu Woo, Seungbin Yoon, HyungGon Shin, Keehoon Kim, Geon Hwee Kim, and Geunbae Lim, "Saline based microfluidic soft pressure sensor utilizing a three-dimensional focused electric field for motion and healthcare monitoring", *Biosensors and Bioelectronics*, vol. 267, no. 116868, January 2025, `https://doi.org/10.1016/j.bios.2024.116868`.

[54] Zhe Su, Jeremy Fishel, Tomonori Yamamoto, and Gerald Loeb, "Use of tactile feedback to control exploratory movements to characterize object compliance", *Frontiers in neurorobotics*, vol. 6, no. 7, July 2012, `https://doi.org/10.3389/fnbot.2012.00007`.

[55] Tianzong Xu, Lijun Li, Yi Wang, Qian Ma, Congying Jia, and Changsheng Shao, "Highly sensitive soft optical fiber tactile sensor", *Optics Express*, vol. 30, no. 19, pp. 34064–34076, September 2022, `https://doi.org/10.1364/OE.467865`.

[56] Yuezhi Cai, Yan Liu, Guangde Li, Qi Qin, Lezhi Pang, Wenhua Ren, Jie Wei, and Muguang Wang, “Reflective tactile sensor assisted by multimode fiber-based optical coupler and fiber specklegram”, *Optics Laser Technology*, vol. 160, no. 109062, May 2023, `https://doi.org/10.1016/j.optlastec.2022.109062`.

[57] Luca Massari, Calogero M. Oddo, Edoardo Sinibaldi, Renaud Detry, Joseph Bowkett, and Kalind C. Carpenter, “Tactile sensing and control of robotic manipulator integrating fiber bragg grating strain-sensor”, *Frontiers in Neurorobotics*, vol. 13, no. 8, 2019, `https://doi.org/10.3389/fnbot.2019.00008`.

[58] Yan Liu, Xiaoli Zou, Qi Qin, Guangde Li, Zhongwei Tan, Muguang Wang, and Fengping Yan, “An optical contact force sensor for tactile sensing based on specklegram detection from concatenated multimode fibers”, *Optics Laser Technology*, vol. 143, no. 107362, November 2021, `https://doi.org/10.1016/j.optlastec.2021.107362`.

[59] H. Maekawa, K. Tanie, and K. Komoriya, “A finger-shaped tactile sensor using an optical waveguide”, in *Proceedings of IEEE Systems Man and Cybernetics Conference - SMC*, Le Touquet, France, 1993, vol. 5, pp. 403–408, IEEE, `https://doi.org/10.1109/ICSMC.1993.390885`.

[60] Rui Li and Edward H. Adelson, “Sensing and recognizing surface textures using a gelsight sensor”, in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, USA, 2013, pp. 1241–1247, IEEE, `https://doi.org/10.1109/CVPR.2013.164`.

[61] K. Kamiyama, K. Vlack, T. Mizota, H. Kajimoto, K. Kawakami, and S. Tachi, “Vision-based sensor for real-time measuring of surface traction fields”, *IEEE Computer Graphics and Applications*, vol. 25, no. 1, pp. 68–75, January 2005, `https://doi.org/10.1109/MCG.2005.27`.

[62] Wael Othman, Zhi-Han A. Lai, Carlos Abril, Juan S. Barajas-Gamboa, Ricard Corcelles, Matthew Kroh, and Mohammad A. Qasaimeh, “Tactile sensing for minimally invasive surgery: Conventional methods and potential emerging tactile technologies”, *Frontiers in Robotics and AI*, vol. 8-2021, January 2022, `https://doi.org/10.3389/frobt.2021.705662`.

[63] Samuel B. Kesner and Robert D. Howe, "Design principles for rapid prototyping forces sensors using 3-d printing", *IEEE/ASME Transactions on Mechatronics*, vol. 16, no. 5, pp. 866–870, July 2011, `https://doi.org/10.1109/TMECH.2011.2160353`.

[64] P. Valdastri, K. Harada, A. Menciassi, L. Beccai, C. Stefanini, M. Fujie, and P. Dario, "Integration of a miniaturised triaxial force sensor in a minimally invasive surgical tool", *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 11, pp. 2397–2400, Novenber 2006, `https://doi.org/10.1109/TBME.2006.883618`.

[65] Nguyen Thanh-Vinh, Nguyen Binh-Khiem, Hidetoshi Takahashi, Kiyoshi Matsumoto, and Isao Shimoyama, "High-sensitivity triaxial tactile sensor with elastic microstructures pressing on piezoresistive cantilevers", *Sensors and Actuators A: Physical*, vol. 215, no. 15, pp. 167–175, August 2014, `https://doi.org/10.1016/j.sna.2013.09.002`.

[66] N B Narayanan, A Bonakdar, J Dargahi, M Packirisamy, and R Bhat, "Design and analysis of a micromachined piezoelectric sensor for measuring the viscoelastic properties of tissues in minimally invasive surgery", *Smart Materials and Structures*, vol. 15, no. 6, pp. 1684, October 2006, `https://doi.org/10.1088/0964-1726/15/6/021`.

[67] Cheng-Hsin Chuang, Tsan-Hsiu Li, I-Chinms Chou, and Ying-Juimr Teng, "Piezoelectric tactile sensor for submucosal tumor detection in endoscopy", *Sensors and Actuators A: Physical*, vol. 244, pp. 299–309, June 2016, `https://doi.org/10.1016/j.sna.2016.04.020`.

[68] Feng Ju, Yaming Wang, Zhao Zhang, Yaoyao Wang, Yahui Yun, Hao Guo, and Bai Chen, "A miniature piezoelectric spiral tactile sensor for tissue hardness palpation with catheter robot in minimally invasive surgery", *Smart Materials and Structures*, vol. 28, no. 2, pp. 25–33, January 2019, `https://doi.org/10.1088/1361-665X/aafc8d`.

[69] Chunyan Li, Pei-Ming Wu, Soohyun Lee, Andrew Gorton, Mark J. Schulz, and Chong H. Ahn, "Flexible dome and bump shape piezoelectric tactile sensors using pvdf-trfe copolymer", *Journal of Microelectromechanical Systems*, vol. 17, no. 2, pp. 334–341, April 2008, `https://doi.org/10.1109/JMEMS.2007.911375`.

[70] Kazuto Takashima, Keisuke Ota, Masaki Yamamoto, Makoto Takenaka, Satoshi Horie, and Kenji Ishida, "Development of catheter-type tactile sensor composed of polyvinylidene fluoride (pvdf) film", *ROBOMECH Journal*, vol. 6, no. 19, pp. 0–11, December 2019, `https://doi.org/10.1186/s40648-019-0147-9`.

[71] Yohan Noh, Hongbin Liu, Sina Sareh, Damith Suresh Chathuranga, Helge Würdemann, Kawal Rhode, and Kaspar Althoefer, "Image-based optical miniaturized three-axis force sensor for cardiac catheterization", *IEEE Sensors Journal*, vol. 16, no. 22, pp. 7924–7932, Novenber 2016, `https://doi.org/10.1109/JSEN.2016.2600671`.

[72] Panagiotis Polygerinos, Tobias Schaeffter, Lakmal Seneviratne, and Kaspar Althoefer, "A fibre-optic catheter-tip force sensor with mri compatibility: A feasibility study", in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, USA, 2009, pp. 1501–1054, IEEE, `https://doi.org/10.1109/IEMBS.2009.5334163`.

[73] Hao Su, Michael Zervas, Gregory A. Cole, Cosme Furlong, and Gregory S. Fischer, "Real-time mri-guided needle placement robot with integrated fiber optic force sensing", in *2011 IEEE International Conference on Robotics and Automation*, Shanghai, China, 2011, pp. 1583–1588, IEEE, `https://doi.org/10.1109/ICRA.2011.5979539`.

[74] R. Ahmadi, J. Dargahi, M. Packirisamy, and R. Cecere, "A new hybrid catheter-tip tactile sensor with relative hardness measuring capability for use in catheter-based heart surgery", in *SENSORS, 2010 IEEE*, Waikoloa, HI, USA, 2010, pp. 1592–1595, IEEE, `https://doi.org/10.1109/ICSENS.2010.5690287`.

[75] Naghmeh Bandari, Roozbeh Ahmadi, Amir Hooshiar, Javad Dargahi, and Muthukumaran Packirisamy, "Hybrid piezoresistive-optical tactile sensor for simultaneous measurement of tissue stiffness and detection of tissue discontinuity in robot-assisted minimally invasive surgery", *Journal of Biomedical Optics*, vol. 22, no. 7, pp. 1–10, July 2017, `https://doi.org/10.1117/1.JBO.22.7.077002`.

[76] Shingo Kawano, Motohiro Kojima, Yoichi Higuchi, Motokazu Sugimoto, Koji Ikeda, Naoki Sakuyama, Shinichiro Takahashi, Ryuichi Hayashi, Atsushi Ochiai, and Norio Saito, "Assessment of elasticity of colorectal cancer tissue, clinical utility, pathological

and phenotypical relevance", *Cancer Science*, vol. 106, no. 9, pp. 1232–1239, June 2015, `https://doi.org/https://doi.org/10.1111/cas.12720`.

[77] Hansoul Kim, Ozdemir Can Kara, and Farshid Alambeigi, "A soft and inflatable vision-based tactile sensor for inspection of constrained and confined spaces", *IEEE Sensors Journal*, vol. 23, no. 23, pp. 29605–29618, December 2023, `https://doi.org/10.1109/JSEN.2023.3324930`.

[78] Brown D. C., "Decentering distortion of lenses", *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.

[79] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart, "A toolbox for easily calibrating omnidirectional cameras", in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006, pp. 5695–5701, IEEE, `https://doi.org/10.1109/IROS.2006.282372`.

[80] Nvidia, "Isaac sim simulator", `https://web.archive.org/web/20251202185725/https://github.com/isaac-sim`.

[81] Nvidia, "Physx5", `https://web.archive.org/web/20251202185848/https://github.com/NVIDIA-Omniverse/PhysX`.

[82] Andrew Nealen, Matthias Müller, Richard Keiser, Eddy Boxerman, and Mark Carlson, "Physically based deformable models in computer graphics", *Computer Graphics Forum*, vol. 25, no. 4, pp. 809–836, December 2006, `https://doi.org/10.1111/j.1467-8659.2006.01000.x`.

[83] Sarah F. Frisken Gibson and Brian Mirtich, "A survey of deformable modeling in computer graphics", Tech. Rep. TR97-19, Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, 1997, `https://www.merl.com/publications/TR97-19/`.

[84] Carmelo Sferrazza and Raffaello D'Andrea, "Sim-to-real for high-resolution optical tactile sensing: From images to three-dimensional contact force distributions", *Soft Robotics*, vol. 9, no. 5, pp. 923–937, October 2021, `https://doi.org/10.1089/soro.2020.0213`.

[85] Kaiming He, X. Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition", in *2016 IEEE Conference on Computer Vision and Pattern*

Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778, IEEE, `https://doi.org/10.1109/CVPR.2016.90`.

[86] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification", in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015, pp. 1026–1034, IEEE, `https://doi.org/10.1109/ICCV.2015.123`.

[87] National Institutes of Health, "Heart library—case model 108 raa with alsca", `https://web.archive.org/web/20251202190113/https://3d.nih.gov/entries/3DPX-016158`.

[88] Louisa O' Neill, Rashed Karim, Rahul K Mukherjee, John Whitaker, Iain Sim, James Harrison, Orod Razeghi, Steven Niederer, Tevfik Ismail, Matthew Wright, Mark D O' Neill, and Steven E Williams, "Pulmonary vein encirclement using an ablation index-guided point-by-point workflow: cardiovascular magnetic resonance assessment of left atrial scar formation", *EP Europace*, vol. 21, no. 12, pp. 1817–1823, December 2019, `https://doi.org/10.1093/europace/euz226`.

[89] Sinuo Zhao, Chi Nguyen, Trung Hoang, Thanh Nho Do, and Hoang-Phuong Phan, "Transparent pneumatic tactile sensors for soft biomedical robotics", *Sensors*, vol. 23, no. 12, pp. 5671, June 2023, `https://doi.org/10.3390/s23125671`.

[90] Mohammad Amin Mirzaee and Ali Sadighi, "Multiphysics simulation and design framework for developing a vision-based tactile sensor with force estimation and slip detection capabilities", *Sensors and Actuators A: Physical*, vol. 378, pp. 115761, November 2024, `https://doi.org/10.1016/j.sna.2024.115761`.

[91] Vijay Kakani, Xuenan Cui, Mingjie Ma, and Hakil Kim, "Vision-based tactile sensor mechanism for the estimation of contact position and force distribution using deep learning", *Sensors*, vol. 21, no. 5, pp. 1920, May 2021, `https://doi.org/10.3390/s21051920`.