Contact Localization of Continuum and Flexible Robot Using Data-driven Approach

Xuan Thao Ha, Di Wu, Chun-Feng Lai, Mouloud Ourak, Gianni Borghesan,
Arianna Menciassi and Emmanuel Vander Poorten

***Abstract*—Continuum robots such as robotic catheters are increasingly being used in minimally invasive surgery. Compliance contributes to enhanced safety during e.g. catheter insertion, however, estimation of contact force and location may help clinicians avoiding exerting excessive force. Ultimately this could lead to faster and safer interventions. Researchers proposed force sensors integrated in the catheter tip in the past. However, such sensors add extra complexity to the catheter design. Also, tip force sensors do not provide insights on forces that act along the catheter length. This paper proposes a data-driven approach for localizing contact forces that appear over the length of the catheter. The proposed approach consists of a collision detection method and a contact localization method. The framework only requires the measurement of the catheter's shape which can be done by an embedded multi-core Fiber Bragg Grating fiber. The method was validated experimentally with a 3D-printed continuum robot with an integrated multi-core fiber. A second contact localization method which is based on identifying the discontinuity in the measured curvature, is also implemented and compared with the proposed method. The static and dynamic experiments show a mean average localization error of 2.3 mm and 4.3 mm which correspond to respectively 3.3% and 6.1% of a 70 mm long flexible robot. These findings demonstrate that the proposed framework outperforms the previous methods and yield promising results. The contact state estimation algorithm can detect collisions in at most approximately $1.08s$.**

I. INTRODUCTION

In the last decade, continuum robots (CRs) have gained popularity in medical robotics thanks to their ability to reach the operational site via torturous lumens, and inherent safety. This is because CRs can adjust shape to adapt to the surrounding anatomy. This feature also helps CRs reach places in the human body that traditional rigid robots cannot access.

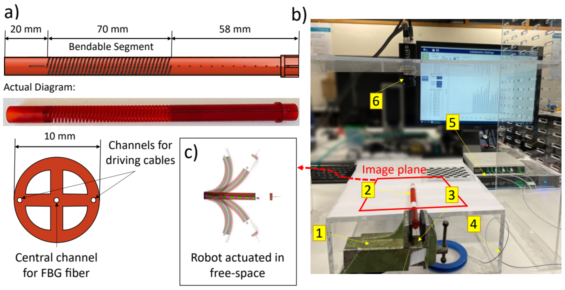
a)
20 mm
70 mm
58 mm
Bendable Segment
Actual Diagram:
10 mm
Channels for driving cables
Central channel for FBG fiber
c)
Robot actuated in free-space
b)
Image plane
1
2
3
4
5
6

Figure 1. a) Design of the CR used in this study and its actual diagram. The top and side view of the robot are shown respectively in the top and bottom figure; b) bench top experiment setup including: 1-mechanical clamp; 2-CR; 3-Nitinol fiber protective tube; 4-multicore FBG fiber; 5-FBG interrogator and 6-overhead camera; c) overlaid images recorded by the overhead camera.

Despite the aforementioned advantages, precise navigation of CRs through a deformable lumen remains a challenging task since the resulting motion is not only determined by the control action but also by the interaction taking place between the CR and the surrounding environment. Additionally, exertion of large forces on soft tissue may cause a variety of complications such as perforation [1]. Researchers developed different controllers to overcome this issue. Controllers based on Cosserat rod modeling were introduced in [2], [3]. However, these controllers require a means to measure the external forces that are applied on the CR. Miniature force sensor that can be embedded in the robot's tip has been proposed [4]. Sterilization and miniaturization problems unfortunately restrict their use in the operating room. Reliable distributed force sensors still did not appear. To avoid extra sensors, researchers proposed estimating the tip force by using a mechanic model and employing knowledge of the applied actuation commands [5]. However, accurate models are complex and difficult to calibrate, and calibration must be executed for each instrument.To overcome this problem, Feng *et al.* developed a machine learning-based method to estimate the tip contact force [6]. Other researchers made use of the shape and the curvature of the robot to estimate forces. Several methods to estimate the shape of the flexible instruments were summarized in [7]. The curvature can e.g. be measured by an external sensor such as a camera [8] or by embedding Fiber Bragg Grating (FBG) sensors. For example, Khan *et al.* estimate the tip force using a strain measuring technique based on FBG [9]. FBG fibers e.g. in combination with a flexure, have been used to produce force sensors [10], [11], [12]. Also, Noh *et al.* developed a miniaturize force sensor based on fiber-optic technology in [13]. However, these approaches do not readily transfer to distributed sensing. They also do not allow detection of forces that occur over the length of the robot. Bajo and Simaan estimate the contact state and contact location based on the static deviation of the robot's pose from its unconstrained kinematic model [14]. Given knowledge of the contact positions, Aloi *et al.* showed how the magnitude and

Manuscript received January 5, 2022; Revised: April 7, 2022; Accepted May 4, 2022. This paper was recommended for publication by Editor P. Valdastri upon evaluation of the Asscociate Editor and Reviewers' comments. This work was supported by the ATLAS project. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 813782. This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017140, the ARTERY project. *(Corresponding author: Di Wu)*

X.T. Ha is with the Department of Mechanical Engineering, KU Leuven, Belgium, and also with The BioRobotics Institute, Scuola Superiore Sant'Anna, Italy (email: xuanthao.ha@kuleuven.be)

D. Wu is with the Department of Mechanical Engineering, KU Leuven, Belgium, and also with the Faculty of 3mE, TU Delft, the Netherlands (email: di.wu@kuleuven.be)

C. Lai is with the Faculty of 3mE, TU Delft, the Netherlands, and also with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, Italy. (email: c.lai-1@tudelft.nl)

M. Ourak and E. Vander Poorten are with the Department of Mechanical Engineering, KU Leuven, Belgium (email: mouloud.ourak@kuleuven.be; emmanuel.vanderpoorten@kuleuven.be)

G. Borghesan is with the Department of Mechanical Engineering, KU Leuven, Belgium, and also with Flanders Make, Belgium (email: gianni.borghesan@kuleuven.be)

A. Menciassi is with The BioRobotics Institute, Scuola Superiore Sant'Anna, Italy (email: arianna.menciassi@santannapisa.it)

Digital Object Identifier (DOI): see top of this page.

direction of the interaction forces can be estimated through numerical finite element modeling (FEM) [15]. Heunis *et al.* [16] presented a method to monitor the distributed contact force of a catheter in real-time under assumption of known contact location. Other researchers developed approaches that jointly estimate the contact force and location applied along the length of the CR. They used hereto the curvature that is calculated from the wavelength shifts measured by a multi-core FBGs [17], [18]. Qiao *et al.* demonstrated in [19] that it is possible to estimate the contact force applied on a flexible instrument by combining the instrument's mechanic model, shape measured from FBG and an Extended Kalman Filter. These methods [17], [18], [19] were verified with a Nitinol rod surrounding an embedded multi-core FBG fiber. However such Nitinol rod is not the best representation of a CR as it lacks actuation.

The accuracy of the curvature-based force estimation approaches greatly depends on the employed contact point localization method. If the contact location is accurately estimated, the magnitude and direction of interaction force can be figured out relatively easy by Cosserat rod model inversion [9]. This aspect forms the starting point of this paper as it concentrates on developing an improved method to accurately localize the contact points. This paper presents a new method to detect collisions with the environment and derives from that a way to localize the contact point along the length of a CR. The proposed method makes only use of the 3D shape as it is being measured by a multi-core FBG fiber. A specific data-driven machine learning (ML) approach, namely an 'AutoEncoder', is used to detect the contact event. This allows a drastic reduction of needed training data and time. In contrast to the method described by Sefati *et al.* in [20], thanks to the use of an AutoEncoder, it suffices to train the model during robot free-space movement. Unlike in [18], which locates the contact point by searching for a discontinuity in the robot's curvature measured by FBGs. In this work, the contact location is estimated by identifying the discontinuity in the tip estimation error. The proposed approach is validated by applying a single contact force in different conditions to a monolythic 3D printed CR equipped with a multi-core FBG fiber inserted in the central working channel. However, with moderate adjustments this method could be extended to localize multiple contact locations. The contact localization method presented in [18] is re-implemented and serves as a baseline against which the newly proposed approach is compared to.

The rest of this paper is organized as follows: Sec. II details the design of the CR and the experimental setup that has been built to evaluate the proposed method. Section III presents the proposed collision detection and contact point localization method. The experimental validation and results are described in Sec. IV. Section V concludes the paper and outlines potential future work directions.

II. DESIGN OF THE CONTINUUM ROBOT

A bidirectional steerable CR was created to mimic a typical steerable catheter. This robot, shown in Fig. 1a, was used for data collection and for verifying the value of the suggested method. The robot is based on a monolythic 3D-printed mechanism called 'HelicoFlex' presented by Culmone *et al.* [21]. The length of its steerable segment is 70 mm. The catheter's outer diameter is 10 mm. There are two steering cables that pass through the proximal rigid segment and the bendable segment to the distal end of the continuum robot. The driving cables are glued to the tip of the CR. When actuators pull the cables the CR will bend sidewards. Note that in this paper, to reduce the complexity of the work, the driving cables are actuated manually. A center channel with an inner diameter of 0.6 mm houses a multicore-FBG fiber. The CR was 3D printed using Perfactory 4 Mini XL's Digital Light Processing technology with R5 epoxy photopolymer resin (EnvisionTec GmbH, Gladbeck, Germany). The precision of the printing layer is 0.03 mm.

To generate ground truth data for training the neural network and validate the algorithm experimentally, a bench top experimental setup has been built around the CR as shown in Fig. 1b. The setup includes three main components: (i) the CR as depicted in Fig. 1a; (ii) an interrogator (FBG-Scan 804, FBGS, Geel, Belgium) to measure the integrated multicore FBG fiber; (iii) an overhead monocular camera (Prosilica, Allied Vision Technology, Germany). The robot is equipped with a 0.25 mm four-core fiber consisting of 32 FBGs from FBGS (Geel, Belgium). The spacing between two gratings is 1 cm as can be seen in Fig. 2b. The fiber is firstly fixed to a protective Nitinol tube by mean of tape at the proximal and distal end. The protective tube is then inserted into the central working channel of the robot. At the proximal and distal end, the protective tube is rigidly fixed to the center channel using epoxy glue. The interrogator enables measuring the wavelength shifts of the FBG sensors when the robot is bent. This can be either due to actuation/pulling of the steering tendons or by external forces applied on the robot's outer surface. Note that the CR is placed on top of a plexiglas box and was constrained by a mechanical clamp at the proximal end. Also note that while the proposed method is general and works in all three dimensions, experiments in 2D have been conducted for sake of readability and verification symplicity. A monocular camera is mounted approximately 30 cm above the CR facing downward to capture the 2D shape of the robot. By using the camera, both the position of the obstacle and of the CR can be identified in the image frame. This then allows calculating accurately the ground truth, namely the actual contact point (if present) between robot and obstacle. This knowledge is then used to validate the method. Figure 1c shows an example of an image recorded by the overhead camera .

During the data gathering, to train the neural network, the wavelength shifts are measured at 200 Hz. However, due to the limitation of the camera which can only stream images at 30 Hz, the wavelength shifts are used only at 30 Hz in the validation experiments. All the sensory data is recorded and synchronized using Robot Operating System (ROS) [22] and then processed by MATLAB (The Mathworks, Massachusetts, US) on a computer with Intel 2.1 GHz core i7 processor and 32 GB of RAM, running Ubuntu 20.04.

III. METHOD

In this paper, discrete curvatures along the length of the CR are measured by distributed multi-core FBG sensors. A method

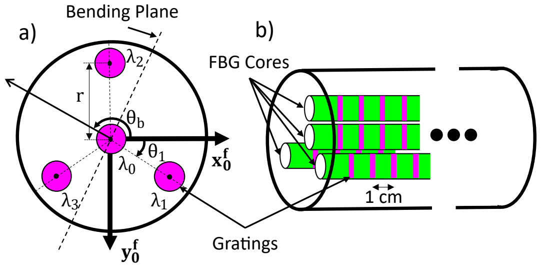

Figure 2. A multi-core FBG fiber with 4 cores and 8 grating sets per core. A cross section and side view of the multi-core fiber can be seen in a) and b), respectively. The angle of bending plane and the angle of the 1st core with respect to the x-axis are denoted as θ_b and θ_1 in a), respectively.

to estimate the curvature κ and the angle of the bending plane θ_b from the measured wavelength shifts of the FBG sensors is briefly described in Sec. III-A. The set of calculated κ's is then used as the input of a data-driven contact detection method (which is detailed in Sec. III-B). In case a collision is detected, the contact point is estimated using the proposed contact localization method. This forms the topic of Sec. III-C. The following assumptions are made in this work:

- the instrument under investigation has a large compliance similar to a catheter or guidewire;
- a single quasi-static external point force is applied. At this stage distributed forces or large dynamics are not taken into account. Nevertheless, the proposed method is assumed to be general and could be extended to estimate a plurality and fast-varying external forces;
- the multi-core fiber is embedded in such a way that disturbances caused by twist are negligible. The reader can refer to Ha *et al.* for a method that reduces the influence of twist [23], after which the here presented method could be applied;
- validation takes place on a planar case, but the method is general in all 3 dimensions (3D).

A. 3D Shape Reconstruction Using Multi-core FBG Fiber

Figure 2 shows the construction of the employed multi-core FBG fiber. The fiber features 4 cores with 8 gratings distributed along each core (leading to a total of 32 gratings). Each grating is a Bragg reflector that reflects a narrow band of the spectrum of incoming light. The rest of the spectrum passes to the next grating or fiber end. The gratings are made in such a way that the spacing and resulting center wavelength of each grating is sufficiently separated from the other. As a result, the reflected light from one grating can be clearly differentiated from light reflected from the other gratings. The Bragg wavelength λ_B is affected by both temperature variation ΔT and mechanical strain ε applied on the optical fiber. The central wavelength of the reflected light of each grating is measured by an interrogator connected to the optical fiber. The measured wavelength shift of the fiber, is then given by:

$$\frac{\lambda_B - \lambda_{B_0}}{\lambda_{B_0}} = \frac{\Delta\lambda}{\lambda_{B_0}} = S_\varepsilon \varepsilon + S_T \Delta T, \tag{1}$$

where λ_{B_0} is the grating's Bragg wavelength in unstrained state, S_ε and S_T are the strain and temperature sensitivity coefficients, respectively. A multi-core fiber typically has a central core that coincides with the neutral axis of the fiber. Gratings in the central core are not sensitive to bending strain, only to axial strain and temperature variations. In case the axial loading is negligible which is often the case, wavelength shifts caused by temperature differences can be calculated. The bending strain on the outer cores can then be determined as:

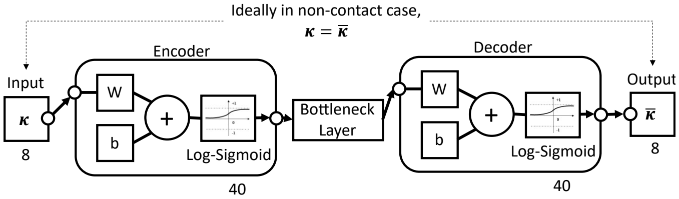
Ideally in non-contact case,
$\boldsymbol{\kappa} = \overline{\boldsymbol{\kappa}}$
Input
$\boldsymbol{\kappa}$
8
Encoder
W
b
+
Log-Sigmoid
40
Bottleneck Layer
Decoder
W
b
+
Log-Sigmoid
40
Output
$\overline{\boldsymbol{\kappa}}$
8

Figure 3. The construction of the built AutoEncoder network that can be used to detect the collision state. The AutoEncoder consists of the encoder, the decoder and the bottleneck layer.

$$\varepsilon_{Bend_{i\in\{2,3,4\}}} = \frac{\Delta\lambda_i}{\lambda_{B_{0i}} S_\varepsilon} - \frac{\Delta\lambda_1}{\lambda_{B_{01}} S_\varepsilon}. \tag{2}$$

The multi-core fiber provides 4 strain measurements $\varepsilon_{i\in\{2,3,4\}}$ at each cross-section where a sensor is present. Given the strain measurements at different locations along the fiber, κ and θ_b can be calculated from:

$$\varepsilon_{Bend_{i\in\{2,3,4\}}} = -\kappa r \sin\left(\theta_b - \frac{3\pi}{2} - \theta_i\right), \tag{3}$$

where r is the distance between the outer cores and the central core; θ_i is the angle of the ith core and θ_b the angle of the bending plane, all expressed with respect to the x-axis (Fig. 2a). A closed-form solution for (3) can be found in [24].

From the κ and θ_b computed at discrete points along the fiber length, the 3D shape of the CR can be reconstructed. The curvature $\kappa(s)$ and torsion $\tau(s)$ profiles can be used to define a continuous space curve as function of arc length s. The rate of change of the angle of the bending plane with respect to s is described as torsion $\tau(s) = \frac{d\theta_b}{ds}$. Curvature and torsion profiles define how the tangent $\mathbf{T}$, normal $\mathbf{N}$, and binormal $\mathbf{B}$ unit vectors change along the length of the curve. The differential Frenet-Serret formula describes this mathematically [25]. The discontinuous curvatures and torsions are then interpolated to enhance shape estimation and provide a quasi-continuous distribution over the arc length of the fiber. The Cartesian position $C(s)$ of each point along the robot's length can then be calculated by integrating:

$$C_s(s) = C_{s,0} + \int_0^l \mathbf{T}(s) ds, \tag{4}$$

where $C_{s,0}$ is the position of the CR's base and l is the robot's length.

B. Detection of Contact State

In this paper, the problem of contact state detection is approached in first instance as a classification problem with two classes: contact and no-contact. The inputs of this classification problem are the discrete curvatures calculated with the method from Sec. III-A, the output is the contact state.

1) Training Data Creation: An unsupervised ML model is trained using only κ at discrete points at each robot configuration while the robot is bending in free-space (no-contact). To generate the training data, the CR is manually actuated by pulling the two driving cables respectively. As the robot is moving the wavelength shifts are recorded to calculate κ at discrete points where the gratings are presented.

2) Data-driven Contact Detection Method: Contact detection algorithms are normally developed based on physical robot models. However, exact physical models are complex to derive. This may lead to inaccurate contact state estimation. A data-driven collision detection method such as presented by Sefati *et al.* [20] could overcome this problem. However, that method would require a very large data-set. Also, this method requires the training data to be generated in both contact and no-contact cases. In the here proposed method, the contact detection problem is formulated as an anomaly detection problem. Anomaly detection can be solved by using an AutoEncoder. Figure 3 shows how an AutoEncoder network consists of an Encoder, Decoder and a bottleneck layer. The Autoencoder can find a reduced dimension representation of the data by focusing on the important characteristics and removing noise and redundancy. During the training phase, only no-contact data are utilized as input to the Encoder. The decoder reconstructs no-contact data from the output of the bottleneck layers. The representation of the no-contact data will be learned by the bottleneck layer since the network will attempt to reduce the difference between the input and output of the AutoEncoder. The contact data is normally different from the no-contact data which makes it difficult for the AutoEncoder to reconstruct the data. Consequently, the reconstruction error will be high. Thanks to this property, the AutoEncoder can detect contact data based on a specified threshold value for the Root Mean Square (RMS) reconstruction error of curvature (E_κ). The advantage of using an AutoEncoder to solve the problem of contact state detection is that only collision-free data (data captured during free-space motion) need to be recorded for training. This forms a significant efficiency enhancement as models that would be trained on contact situations would need to see every possible contact situation to be able to accurately predict a contact. Let $\{\kappa\}_{t=1}^{N}$ be a series of observations with $\kappa \in \mathbb{R}^m$ representing the observation at time t and m the number of gratings. Given the training set $\{\kappa\}_{t=1}^{N}$, an AutoEncoder network can be trained to detect the contact state. To determine whether a collision occurs, an empirically-derived threshold value can be imposed on E_κ. Figure 5a summarizes the data-driven contact detection method.

C. Contact Location Estimation

Upon detection of contact, the next step is to find the contact location. First, an earlier method [17], [18], that searches a discontinuity in the curvature (DiC), is explained (SubSec.III-C1). Then, SubSec.III-C2 presents a new method that analyses the discontinuity in the tip estimation error (DiTE).

1) discontinuity in curvature (DiC): Let us consider the situation in which only point forces are applied to the robot. In such case, according to the Cosserat rod's constitutive equations [26], the internal moment $\mathbf{m}$ of each point along the robot can be calculated from the angular strain vector $\mathbf{u}$ and the bending stiffness matrix $\mathbf{K}_\mathrm{u}$. The vector $\mathbf{u}$ can then be obtained from the calculated curvature κ and θ_b as:

$$\mathbf{u} = \begin{bmatrix} \kappa \cos\theta_b \\ \kappa \sin\theta_b \\ 0 \end{bmatrix}. \tag{5}$$

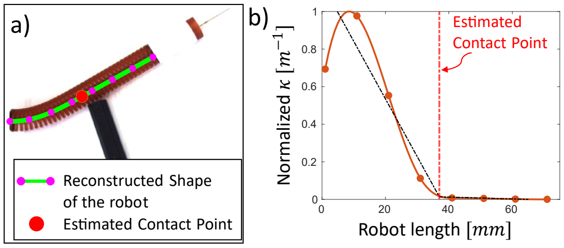

Figure 4. a) The 3D shape of the robot was reconstructed and transformed to the image plane (pre-registration step is explained in section IV-B). The contact point estimated by finding the discontinuity in measured curvature is marked in red. b) The measured curvature was normalized and a piece-wise linear approximation method was used to identify the discontinuity point.

It is assumed in this work that the CR is solely subjected to a distributed moment generated by pulling the driven cables and the compression stiffness is infinite. Thus, the internal moment $\mathbf{m}$ should vary continuously. However, when external point forces are exerted, the internal moment will be disrupted. Based on this observation, the contact locations can be found at the points where κ undergoes a discontinuity. Figure 4 shows this principle. This approach has been described and verified in [17], [18] on a rod equipped with FBG fiber in a center channel. The contact localization method based on identifying the discontinuity in κ has been re-implemented for comparison on the bench top setup. It is worth mentioning that this is the first time this DiC-based contact localization algorithm is applied on a steerable instrument. To find the contact point, a piece-wise linear approximation algorithm is used to find the discontinuity (break point) in the measured curvature. The piece-wise linear approximation algorithm starts by assuming each point along the robot length is a break point. The set of κ is split into two subsets at the break point. Each subset will be fitted with a linear model using *fitlm* function provided by MATLAB. The piece-wise linear fitting error will be then calculated. The break point with the lowest fitting error will be considered as the actual contact point. To improve the accuracy, the set of measured curvature is normalized before applying the piece-wise linear approximation algorithm.

The experiment results show that the contact estimation's accuracy of this method is not consistent. While the CR is only bent due to external force (no actuation), the accuracy of the estimation is equivalent to the results given in [17], [18]. However, the accuracy degrades when the CR simultaneously experiences external forces and a distributed moment by actuating via the steering cable. The experiments and the estimation accuracy of the DiC-based contact localization approach are detailed in Sec. IV.

2) discontinuity in tip estimation error (DiTE): To improve the estimation accuracy in case the drive cables are actuated, we propose a method that can identify the contact point by finding the discontinuity in the error between the measured robot's tip position and the expected estimated tip position. A free-space curvature model of the CR is first approximated using a ML model. This free-space curvature model allows estimating the curvature and the angle of bending plane at all discrete points where gratings are present while knowing κ and θ_b at a random grating along the CR length. In this work, a k-Nearest Neighbors (kNN) model is used to fit the free-space curvature model. kNN is one of the simplest ML algorithms, but can be effective for classification based on supervised learning. The main disadvantages are the limited efficiency and reliance on the selected *k*-value. Since this paper is used to demonstrate the proof of concept of the proposed method, kNN is deemed sufficient. More advanced data-driven free-space curvature models can be implemented in the future to improve accuracy and processing time.

The data collected during training of the AutoEncoder is conveniently re-used to train a kNN model. The model takes 3 inputs: κ, θ_b and i where κ and θ_b are curvature and angle of bending plane calculated at i^{th} grating by the measured wavelength shift ($i = 1 \ldots m$), respectively. The output of the kNN model are the estimated curvatures $\tilde{\kappa}$ and bending plane angles $\tilde{\theta}_b$ of all locations where FBGs are located. We define the estimated curvature and bending plane angle of the jth grating ($j = 1 \ldots m$) by using information from the ith grating as $\tilde{\kappa}_{i,j}$ and $\tilde{\theta}_{b_{i,j}}$, respectively. The general idea is that for motion in free space our robot has excellent repeatability which means that for a given tension to each cable the robot will always take on a very similar shape in space. This shape will correspond to a given κ and θ_b at a certain grating i and also given κ and θ_b at the other gratings. Since the network is trained using data generated in free-space motion, the kNN model learns the relation between the values (including κ and θ_b) at grating i and those at the other gratings, such that from a single pair of values at i all the other can be derived (provided that the robot is in free-space). If contact appears, there will be differences between estimated and measured values from the first grating to the grating where the external force is applied.

Rather than comparing the angle of the bending plane and curvatures directly, it is found convenient to investigate the effect on the estimated robot tip position. For using measurements up to the ith grating $\tilde{C}_{s_i}(l)$ is calculated by integrating $[\kappa_1, \ldots, \kappa_i, \tilde{\kappa}_{i,i+1}, \ldots, \tilde{\kappa}_{i,m}]$ and $[\theta_{b_1}, \ldots, \theta_{b_i}, \tilde{\theta}_{b_{i,i+1}}, \ldots, \tilde{\theta}_{b_{i,m}}]$ with the method of section III-A. Similar the actual tip position $C_s(l)$, estimated by sensor measurements only, can be derived by integrating κ_i and θ_{b_i} from $i = 1, \ldots, m$.

The error in tip position estimation of ith grating E_{Tip_i} is then computed as the Euclidean distance between $\tilde{C}_{s_i}(l)$ and $C_s(l)$. At each time step, E_{Tip_i} is calculated for each grating. Assuming that an external point force is acting at the qth grating, as shown in Fig. 5b one can expect to see the error in the tip estimation of all gratings before the qth grating ($E_{Tip_{i=1,\ldots,q-1}}$) to be significantly larger than 0 while $E_{Tip_{i=q+1,\ldots,m}}$ associated to the remaining gratings will be close to 0. This is because the applied force will not affect the curvature profile of the more distal robot segment (where the qth to m^{th} gratings are present). For this reason, the free-space curvature model is still valid for these gratings, resulting in a smaller error in tip estimation. A piece-wise linear approximation algorithm explained in SubSec.III-C1, is used to find the break point in the error of tip estimation. The found break point will be considered as the contact location. Since the contact can appear at any arbitrary point along the robot length, the errors in tip estimation summarized as $\mathbf{E_{Tip}} = [E_{Tip_1} \ldots E_{Tip_l}]$ are interpolated to improve the resolution of the algorithm. $\mathbf{E_{Tip}}$ is normalized before the piece-wise linear approximation method is applied. The proposed framework to estimate the contact location in case the robot is actuated is shown in Fig. 5b.

The contact detection algorithm is used to identify whether the robot is touching an obstacle. In the event of a collision, the contact localization method then localizes the point of contact.

IV. EXPERIMENTAL VALIDATION

A. Experiment Design

To evaluate the proposed framework, a training data set is first recorded. The robot is actuated in free-space by manually pulling the driven cables to cover its workspace for approximately 2 minutes. The measured wavelength is used to calculate the curvature and angle of bending plane of points where gratings are etched. This training data set is used to train the contact detection model (AutoEncoder network) and the free-space curvature model (kNN). The AutoEncoder hyper-parameters namely the hidden size and max epochs are tuned by using randomized k-fold cross-validation on the training data set. An initial k-fold ($k = 4$) validation took place on hidden sizes of $\{20, 30, 40, 50\}$ and max epochs $\{600, 800, 1200\}$. The performance results on the training data set with 24490 samples with different combinations of hidden size and max epochs is summarized in Table I.

The optimal hyper-parameters were selected to minimize the k-fold mean RMS validation error of performance over all combinations of training and testing sets in a k-fold split. The three sets of hyper-parameters shown in bold in Table I gives comparable validation RMS errors. The small difference in validation error is negligible. The set of parameters with smaller number of epoch and hidden size should be chosen to reduce overfitting and computational time of the network. As such 40 and 1200 were chosen as optimal values for max epoch and hidden size, respectively. Figure 3 shows the construction of the built AutoEncoder. The free-space curvature model is fitted by a kNN model with k value and distance metric set to 1 and Euclidean distance, respectively. Since κ and $\boldsymbol{\theta_b}$ are not on the same scale, these two sets are normalized before applying the kNN method to ensure that all of the features are treated equally when calculating the distance.

The performance of the method is then evaluated on unseen data whereby the CR moves in free-space as well as constraint by the environment. The test set is produced by manually pulling the steer cables and bending the robot in different shapes until the robot collides with obstacles that are placed at 10 different known locations on each side of the CR. Beside the static experiment, a dynamic experiment was also conducted by continuously bending the robot. Bending took place in two directions and contacts where established with obstacles at different points along the length of the robot to show the robustness of the proposed approach (see attached video).

B. Ground Truth Generation

Color thresholding of the image of the overhead camera is first used to segment the CR and the obstacle contour. To reduce background noise and enable strong linked area recognition, an erosion morphological operation is applied on the binary images, followed by a dilation, both with small kernel. The collision is detected when the robot's contour and the obstacle form a connected region.

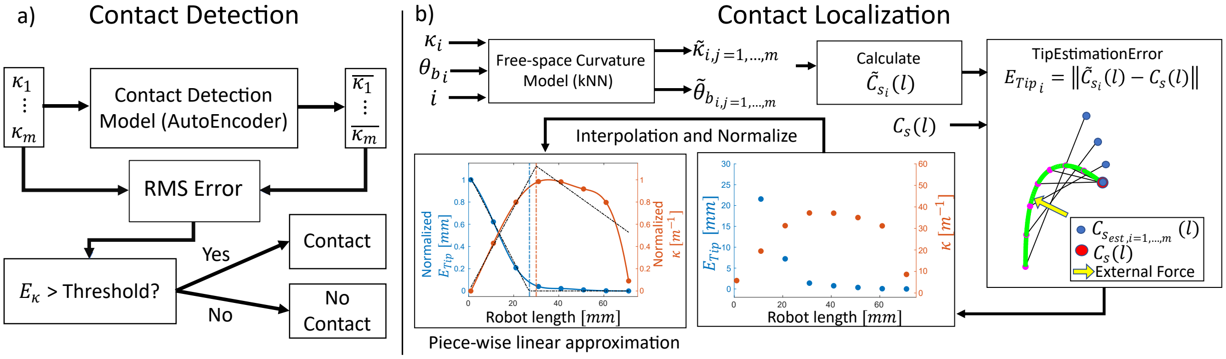
a) Contact Detection
κ_1 ⋮ κ_m
Contact Detection Model (AutoEncoder)
$\bar{\kappa}_1$ ⋮ $\bar{\kappa}_m$
RMS Error
E_κ > Threshold?
Yes
No
Contact
No Contact
b) Contact Localization
κ_i
θ_{b_i}
i
Free-space Curvature Model (kNN)
$\tilde{\kappa}_{i,j=1,\dots,m}$
$\tilde{\theta}_{b_{i,j=1,\dots,m}}$
Calculate $\tilde{C}_{s_i}(l)$
TipEstimationError
$E_{Tip_i} = \|\tilde{C}_{s_i}(l) - C_s(l)\|$
$C_s(l)$
Interpolation and Normalize
Normalized E_{Tip} $[mm]$
Normalized κ $[m^{-1}]$
Robot length $[mm]$
E_{Tip} $[mm]$
κ $[m^{-1}]$
Robot length $[mm]$
$C_{s_{est,i=1,\dots,m}}(l)$
$C_s(l)$
External Force
Piece-wise linear approximation

Figure 5. The proposed framework includes a) a method to detect collision and b) a method to localize the contact point. a) To detect collision state, the RMS reconstructed error is first calculated. An threshold value determined experimentally is used to check whether contact appears. b) The tip estimation error of each gratings are first calculated in contact localization method. After that, interpolation is applied on the tip estimation error to increase the resolution. The interpolated tip estimation error is normalized before piecewise-linear approximation algorithm is applied to identify the location of the contact point. The blue and red dashed lines in b) show the contact location estimated by DiTE-based and DiC-based method, respectively.

Table I
HYPERPARAMETERS FOR THE AUTOENCODER. THREE SETS OF PARAMETERS IN BOLD GIVE COMPARABLE VALIDATION RMS ERRORS. THE OPTIMAL HYPER-PARAMETERS OF 40 AND 1200 FOR MAX EPOCH AND HIDDEN SIZE WERE CHOSEN, RESPECTIVELY.

Max Epoch	800				**1200**				**1400**			
Hidden Size	20	30	40	50	20	30	**40**	50	20	30	**40**	**50**
Mean RMS Validation Error	0.63	0.56	0.59	0.60	0.54	0.50	**0.44**	0.46	0.62	0.50	**0.43**	**0.42**
Training Time (s)	23.71	31.22	31.22	38.00	39.00	36.51	**57.83**	70.71	74.31	91.29	**101.91**	**111.45**

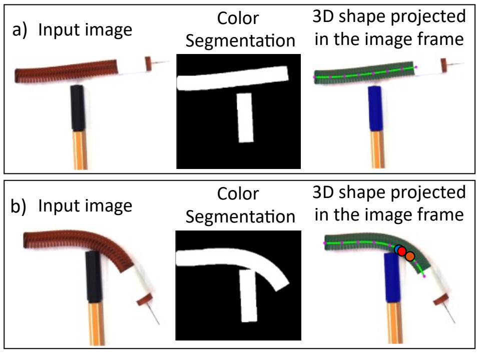
a) Input image
Color Segmentation
3D shape projected in the image frame
b) Input image
Color Segmentation
3D shape projected in the image frame

Figure 6. An example of ground truth generation process in the case where the catheter is moving in free-space (a) and in the case of a collision upon bending to the right (b). The estimated contact points respectively by DiC and DiTE are marked by an orange and blue marker in (b), whereas the ground truth contact location is shown by a red circle.

Additional steps are required to identify the ground truth contact location. Firstly, the 3D reconstructed shape coordinate frame (FBG coordinate frame) is registered with the image coordinate frame. For this pre-registration step, the robot is bent in free-space while the image from the overhead camera and the wavelength shifts are recorded. The contour of the robot is then segmented and a skeletonisation algorithm [27] is applied on the segmented contour to find the robot's center-line in 2D. The 3D shapes of the robot were reconstructed by using the measured wavelength shifts. The set of points which contains all 3D shapes (reconstructed by the measured wavelength shifts) and 2D shapes (recognized in the 2D image) of the robot are defined as $P^{FBG}_{x\times3}$ and $P^{img}_{w\times2}$, respectively where x and w are the number of points in each point cloud. Since the robot is moving on top of a plexiglas plate that is positioned parallel to the image plane, the 2D center-line of the robot can be transformed to a 3D center-line by adding a fixed z-value to $P^{img}_{w\times2}$ (i.e. $z = 0$).

Finally, the transformation matrix T^{FBG}_{img} that is used to transform the FBG coordinate frame to the image coordinate frame is found by using an iterative closest point algorithm [28] on the two sets of point cloud $scale \times P^{FBG}_{x\times3}$ and $[P^{img}_{w\times2}\ 0_{w\times1}]$ where $scale$ is the camera scale factor calibrated by observing a known-size chess board. During the experiment, the 3D reconstructed shape is transformed to the 2D image coordinate frame at each time step by multiplying T^{FBG}_{img} with $scale \times P^{FBG}_{x\times3}$ followed by truncating the z-values. The closest pair of points between the reconstructed shape in the image frame and the obstacle contour was then found. Points that belong to both the reconstructed shape and the pair of closest points is identified as the contact point. Examples of the ground truth generation process in case of no-contact and contact are shown in Fig. 6a and Fig. 6b, respectively. The estimated contact point by the DiC-based and DiTE-based methods and the ground truth contact location are marked with orange, blue and red circles in Fig. 6b, respectively.

C. Results and Discussion

Figure 7a and 7b show an example of experimental data demonstrating the proposed collision detection and contact localization methods, respectively. The continuous black line (referenced to the left y-axis) in Fig. 7a corresponds to the RMS reconstructed error E_κ. To improve the stability of the contact detection algorithm, a moving average filter with 10 samples was applied to the calculated E_κ. The threshold value to detect if contact exists is defined empirically and is set at 3.7. By lowering the threshold value, the contact can be detected more quickly. However, this also increases the chance of false positives in contact state estimation. The contact detection is activated only when E_κ is larger than the predefined threshold value for a period longer than $0.3s$. Also, this threshold window time was set empirically and helped to reduce false positive detections, hence effectively improving the algorithm's robustness. For these choices, Fig. 7a shows that the collision is detected at $t = 5.8$s, thus approximately 1.43s after the contact was established. The contact points

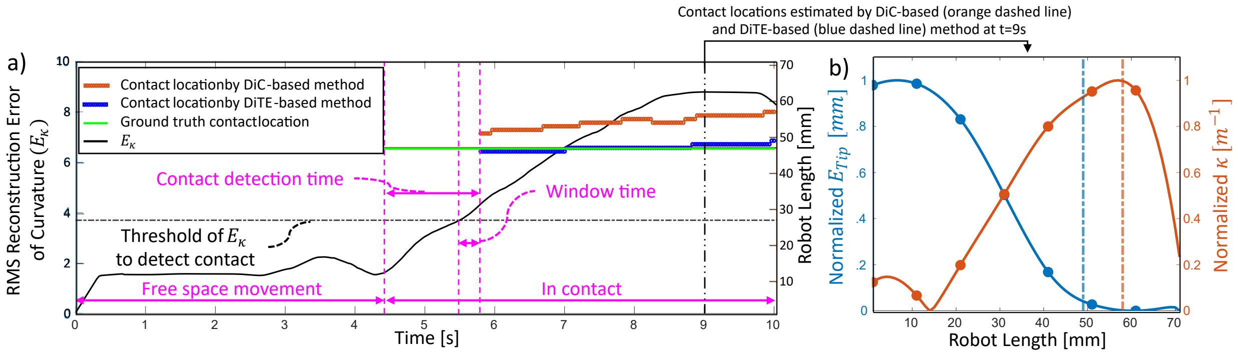

Figure 7. Panel (a) depicts the estimated contact location by the DiC-based (red) and DiTE-based (blue) methods together with the ground truth (green) over time. These locations are referenced to the left y-axis. Panel (b) shows the piece-wise linear approximation when the catheter is in contact. The blue vertical dashed line shows the discontinuity in the calculated normalized E_{Tip} while the red vertical dashed line shows that of the measured normalized curvature κ.

Table II

EXPERIMENTAL RESULTS OF THE STATIC EXPERIMENT. THE CR IS BENT TO BE IN CONTACT WITH THE OBSTACLE AT DIFFERENT LOCATIONS.

		Left Bending					Right Bending				
Ground truth location (mm)		29	35	40	47	57	28	38	48	52	59
Time to detect (s)		1.53	0.77	0.90	0.73	0.87	1.50	0.50	1.43	1.30	1.30
Estimated contact location by	**DiC**	33	45	44	50	56	33	41	54	57	61
	DiTE	28	37	39	46	52	30	38	47	48	53
	Hybrid	28	37	39	46	56	30	38	47	57	61
Standard deviation	**DiC**	1.04	5.71	2.20	6.19	8.27	0.99	0.51	1.62	0.89	0.64
	DiTE	0.20	2.83	1.34	1.59	3.02	0.50	0.85	0.81	0.41	0.95
	Hybrid	0.20	2.83	1.34	1.59	8.27	0.50	0.85	0.81	0.89	0.64

estimated by calculating the DiC and the DiTE are plotted in red and blue in Fig. 7a, respectively. The ground truth contact location is shown in red. The contact localization algorithm only operates when a collision is detected. Figure 7b shows the calculated normalized tip estimation error and normalized curvature at $t = 8$s. The blue and red dashed line in Fig. 7b show the output of the contact localization algorithm of the DiC and DiTE-based methods, respectively.

Table II summarizes the experimental results for 10 static experiments. The average collision detection time was 1.08s. The results indicate that contacts closer to the middle of the robot allow for faster contact detection. By lowering the detection threshold and reducing the window time, collisions can be detected faster. However, this will increase the chance of false positive detections. Therefore researchers ought to trade-off these two parameters based on their application. It is worth noting that the clinician's reaction time during minimally invasive procedures is about 300-400 ms [29], [30]. Although the proposed method is a bit more than twice of this reaction time, it must be stressed that the haptic feedback that clinicians receive now is non-discriminative. Clinicians have poor knowledge of where obstacles and forces are located. Further, given the low speeds (less than 5 mm/s in general [31], [32]) aside from a reduction of mental load, one can also expect that tissue damage can be reduced significantly. The performance of the DiTE-based method tends to reduce when the contact appears at the distal end. The proposed DiTE approach estimates the contact point by identifying the discontinuity in the calculated E_{Tip}. However, the maximal value of E_{Tip} that can be obtained by each point along the length of the CR decreases from the base to the tip. This is caused by the fact that the closer the point to the tip, the fewer estimates of $\tilde{\kappa}$ and $\tilde{\theta}$ are used to estimate the tip position. Hence E_{Tip} will vary less for distal points. The method will thus be less sensitive for those contacts. Despite the fact that the accuracy of DiC-based algorithm increases when external forces are closer to the tip, the DiTE-based on average outperforms the DiC-based contact localization method. To compare the two methods, the mean average percentage error of each method was calculated by dividing the mean average contact estimation error of the ten experiments by the length of the flexible robot section. The mean average percentage contact estimation error from DiC and DiTE-based methods are 6.1% (4.3 mm) and 3.3% (2.3 mm) in the static experiments, respectively. The DiTE-based method also yields a lower mean standard deviation, 1.3 mm compared to 2.8 mm of the DiC-based method. The contact detection processing time is 0.01s while that of contact localization is 0.03s. Further improvements in processing time can be achieved by replacing the kNN-based free-space curvature model by a neural network-based model. A hybrid approach could be implemented to make use of the advantages of both methods by running both estimators in parallel and switching between the DiTE and the DiC methods in terms of the contact location. The described hybrid method features a switching mechanism, in which the estimated location of DiTE method is selected when DiTE-based estimated contact location is in the proximal part of the CR. Otherwise, the estimated contact location of DiC-based method is used. The proximal part of the CR is defined by a percentage value α and the length of the proximal part is calculated as $\alpha\times$ CR's length. The experimental results of the static experiment (see Table II row "hybrid") showed that the hybrid method was more effective than one method alone with the mean average percentage contact estimation error of 2.2% (1.6 mm). In this static experiment, α is set at 68%. In the dynamic experiment, the CR was brought into contact with an obstacle at 11 different locations. The mean average percentage contact estimation errors were found to be 5.4% (3.8 mm) and 8.4% (5.9 mm) for DiTE and DiC-based contact localization methods, respectively. The DiTE-based approach also showed a smaller standard deviation error of 2.7 mm, whereas the DiC- based method showed a standard deviation error of 3.7 mm. The experimental results of the dynamic experiment can be seen in the attached video.

If multiple contacts arise, it is believed that the here presented approach could be extended as explained next. The proposed approach is able to localize a single contact force applied to the robot. In case of multiple forces acting at different contact locations, the proposed approach can serve as a starting step. Previous curvature-based force sensing methods were only evaluated on a passive Nitinol rod with an embedded multi-core FBG [17], [18], [19]. These approaches are not able to account for the distributed forces generated by actuation. The DiTE-based contact localization method is able to identify the last contact force (the closest contact force to the tip) in case of multiple contact forces. Then, one can understand that the curvature of the length of the robot from the last contact location to the tip is caused by the distributed forces induced by the actuation. By using this curvature information, the forces caused by the actuators could be calculated. Therefore in principle the effect of the actuation on the CR could be compensated for and the previous curvature-based force sensing methods [17], [18], [19] can then be applied to derive the other contact locations. After localizing the contact forces, the magnitude of these forces could then be calculated by inversion of the Cosserat rod model [9] or through FEM [15].

V. CONCLUSIONS

This paper presents a new data-driven approach that allows estimating collision state and contact location solely relying on information from FBG sensors. The approach has been empirically validated with a 3D-printed cable-driven robot and yields promising results. The approach can detect the collisions in approximately $1.08s$ and can estimates contact locations with a mean average percentage error of 3.3% in the static experiment. The mean average percentage error of the dynamic experiment using the proposed DiTE approach is 6.1%. In our proposed approach, no assumption was needed regarding the relative position of the multi-core fiber and the robot's centerline. As a result, it is expected that this method will work with multi-core FBG fibers in off-center channels or multiple single-core fibers routed at the circumference, freeing up the central channel for more critical tasks. In the future, the proposed method will be evaluated while the robot is operating in 3D. In this case, multi-view 3D computer vision method can be used to generate the ground truth.

REFERENCES

[1] J. S. Balami, P. M. White, P. J. McMeekin *et al.*, "Complications of endovascular treatment for acute ischemic stroke: prevention and management," *Intl. J. of Stroke*, vol. 13, no. 4, pp. 348–361, 2018.

[2] G. Smoljkic, G. Borghesan, D. Reynaerts *et al.*, "Constraint-based interaction control of robots featuring large compliance and deformation," *IEEE Trans. on Robot.*, vol. 31, no. 5, pp. 1252–1260, 2015.

[3] P. E. Dupont *et al.*, "Design and control of concentric-tube robots," *IEEE Trans. on Robot.*, vol. 26, no. 2, pp. 209–225, 2009.

[4] P. Valdastri, K. Harada, A. Menciassi *et al.*, "Integration of a miniaturised triaxial force sensor in a minimally invasive surgical tool," *IEEE trans. on Biomed. Eng.*, vol. 53, no. 11, pp. 2397–2400, 2006.

[5] C. B. Black, J. Till, and D. C. Rucker, "Parallel continuum robots: Modeling, analysis, and actuation-based force sensing," *IEEE Transactions on Robotics*, vol. 34, no. 1, pp. 29–47, 2017.

[6] F. Feng, W. Hong, and L. Xie, "A learning-based tip contact force estimation method for tendon-driven continuum manipulator," *Scientific Reports*, vol. 11, no. 1, pp. 1–11, 2021.

[7] S. K. Sahu, C. Sozer, B. Rosa *et al.*, "Shape reconstruction processes for interventional application devices: State of the art, progress, and future directions," *Frontiers in Robotics and AI*, vol. 8, 2021.

[8] J. Back, T. Manwell *et al.*, "Catheter contact force estimation from shape detection using a real-time cosserat rod model," in *2015 IEEE/RSJ Intl. Conf. on Intell. Robots and Syst. (IROS)*. IEEE, 2015, pp. 2037–2042.

[9] F. Khan, R. J. Roesthuis, and S. Misra, "Force sensing in continuum manipulators using fiber bragg grating sensors," in *2017 IEEE/RSJ Intl. Conf. on Intell. Robots and Syst. (IROS)*. IEEE, 2017, pp. 2531–2536.

[10] T. Li, C. Shi, and H. Ren, "Three-dimensional catheter distal force sensing for cardiac ablation based on fiber bragg grating," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 5, pp. 2316–2327, 2018.

[11] A. Gao, N. Liu, H. Zhang, Z. Wu, and G.-Z. Yang, "Spiral fbg sensors-based contact detection for confocal laser endomicroscopy," *Biosensors and Bioelectronics*, vol. 170, p. 112653, 2020.

[12] P. Polygerinos, L. D. Seneviratne, R. Razavi *et al.*, "Triaxial catheter-tip force sensor for mri-guided cardiac procedures," *IEEE/ASME Trans. on mechatronics*, vol. 18, no. 1, pp. 386–396, 2012.

[13] Y. Noh, H. Liu, S. Sareh *et al.*, "Image-based optical miniaturized three-axis force sensor for cardiac catheterization," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 7924–7932, 2016.

[14] A. Bajo and N. Simaan, "Finding lost wrenches: Using continuum robots for contact detection and estimation of contact location," in *Intl. Conf. on Robots. and Autom.* IEEE, 2010, pp. 3666–3673.

[15] V. A. Aloi and D. C. Rucker, "Estimating loads along elastic rods," in *Intl. Conf. on Robots. and Autom. (ICRA)*. IEEE, 2019, pp. 2867–2873.

[16] C. M. Heunis *et al.*, "Reconstructing endovascular catheter interaction forces in 3d using multicore optical shape sensors," in *IEEE/RSJ Int. Conf. on Intell. Robot. and Syst. (IROS)*. IEEE, 2019, pp. 5419–5425.

[17] O. Al-Ahmad *et al.*, "Fbg-based estimation of external forces along flexible instrument bodies," *Frontiers in Robotics and AI*, vol. 8, 2021.

[18] Q. Qiao, D. Willems, G. Borghesan *et al.*, "Estimating and localizing external forces applied on flexible instruments by shape sensing," in *Intl. Conf. on Advanced Robots. (ICAR)*. IEEE, 2019, pp. 227–233.

[19] Q. Qiao *et al.*, "Force from shape—estimating the location and magnitude of the external force on flexible instruments," *IEEE Trans. on Robots.*, vol. 37, no. 5, pp. 1826–1833, 2021.

[20] S. Sefati *et al.*, "Learning to detect collisions for continuum manipulators without a prior model," in *Intl. Conf. on Med. Imag. Comput. and Computer-Assisted Intervention*. Springer, 2019, pp. 182–190.

[21] C. Culmone *et al.*, "Exploring non-assembly 3d printing for novel compliant surgical devices," *Plos one*, vol. 15, no. 5, p. e0232952, 2020.

[22] M. Quigley, K. Conley, B. Gerkey *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[23] X. T. Ha *et al.*, "Robust catheter tracking by fusing electromagnetic tracking, fiber bragg grating and sparse fluoroscopic images," *IEEE Sens. Jour.*, vol. 21, no. 20, pp. 23 422–23 434, 2021.

[24] J. P. Moore and M. D. Rogge, "Shape sensing using multi-core fiber optic cable and parametric curve solutions," *Optics express*, vol. 20, no. 3, pp. 2967–2973, 2012.

[25] T. F. Banchoff and S. T. Lovett, *Differential geometry of curves and surfaces*. AK Peters/CRC Press, 2010.

[26] M. B. Rubin, *Cosserat theories: shells, rods and points*. Springer Science & Business Media, 2013, vol. 79.

[27] T.-C. Lee, R. L. Kashyap, and C.-N. Chu, "Building skeleton models via 3-d medial surface axis thinning algorithms," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, pp. 462–478, 1994.

[28] S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," in *Proceedings third international conference on 3-D digital imaging and modeling*. IEEE, 2001, pp. 145–152.

[29] M. Seibold, S. Maurer, A. Hoch *et al.*, "Real-time acoustic sensing and artificial intelligence for error prevention in orthopedic surgery," *Scientific Reports*, vol. 11, no. 1, pp. 1–11, 2021.

[30] B. Zheng, Z. Janmohamed, and C. MacKenzie, "Reaction times and the decision-making process in endoscopic surgery," *Surgical Endoscopy and Other Interventional Techniques*, vol. 17, no. 9, pp. 1475–1480, 2003.

[31] K. W. Eichhorn, R. Westphal *et al.*, "Robot-assisted endoscope guidance versus manual endoscope guidance in functional endonasal sinus surgery (fess)," *Acta oto-laryngologica*, vol. 137, no. 10, pp. 1090–1095, 2017.

[32] Z. Zhang, J. Dequidt, J. Back, H. Liu, and C. Duriez, "Motion control of cable-driven continuum catheter robot through contacts," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1852–1859, 2019.