Vision-based Tip Force Estimation on a Soft Continuum Robot

Xingyu Chen[2], Jialei Shi[1], Helge Wurdemann[1], and Thomas George Thuruthel [2]

***Abstract*— Soft continuum robots, fabricated from elastomeric materials, offer unparalleled flexibility and adaptability, making them ideal for applications such as minimally invasive surgery and inspections in constrained environments. With the miniaturization of imaging technologies and the development of novel control algorithms, these devices provide exceptional opportunities to visualize the internal structures of the human body. However, there are still challenges in accurately estimating external forces applied to these systems using current technologies. Adding additional sensors is challenging without compromising the softness of the device. This work presents a visual deformation-based force sensing framework for soft continuum robots. The core idea behind this work is that point loads lead to unique deformation profiles in an actuated soft-bodied robot. We introduce a Convolutional Neural Network-based tip force estimation method that utilizes arbitrarily placed camera images and actuation inputs to predict applied tip forces. Experimental validation was performed using the STIFF-FLOP robot, a pneumatically actuated soft robot developed for minimally invasive surgery. Our vision-based force estimation model demonstrated a sensing precision of 0.05 N in the XY plane during testing, with data collection and training taking only 70 minutes.**

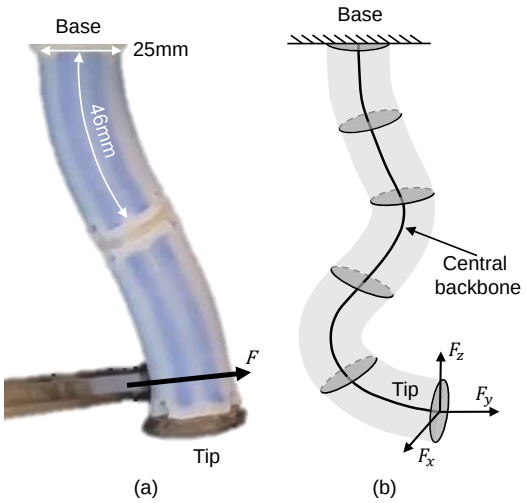

Fig. 1. Vision-based external force estimation. A) The pneumatically actuated soft robot (STIFF-FLOP) used for this study. B) Given an actuator input and external force input, the robot reaches a unique configuration statically, which, if observable, can be used to solve the inverse problem.

I. INTRODUCTION

With their unique flexibility and adaptability, continuum robots have found applications ranging from minimally invasive surgery to inspections in constrained environments [1], [2], [3]. A soft continuum robot epitomizes a particular design within continuum robots, constructed entirely from elastomeric materials, achieving shape control via pneumatic or hydraulic chambers in its elongated components [4]. While their remarkable dexterity makes them highly sought after for surgical robots, they also present significant challenges in external force sensing, a critical aspect for control and safety [5], [6]. This challenge has spurred numerous research efforts aimed at developing specialized force estimation methods tailored for continuum robots.

Model-based techniques, grounded in kinematics and statics, have been devised using mathematical models of their motions to estimate external forces from observed states or deformations [7], [8], [9]. A common model used for this assumes that the deformation of the soft robot to an arc of constant curvature [10]. Xu et al.[11] developed a force-sensing strategy by crafting a kinetostatic model using the constant curvature model and backbone tracking of a continuum robot. Rucker et al. extended this approach using only end-effector poses and an extended Kalman Filter [12]. In 2015, Khoshnam et al. proposed a kinematic model, using local shaft curvature and images to measure robot shape and forces[13]. To better simplify the models, Hasanzadeh et al. [14] proposed a low-dimensional quasi-static model adept at solving the inverse problem, while Bajo et al. [15] proposed a hybrid motion/force controller using similar constant curvature based kinetostatic models. While model-based approaches have displayed potential for these intrinsic force estimation, their usage remains restricted to simple robotic systems, lacks generalizability, and necessitates an arduous calibration process.

Recent years have seen a surge in studies resorting to direct force sensing solutions. For instance, some harness fibre Bragg grating (FBG) sensors for shape or force detection, though they are temperature sensitivity and adds additional complexity to the design [16], [17], [18]. Optical transducers have also been used to develop accurate force sensors [19]. Truby et al.[20] developed a sensing skin that can be installed on the body surface of continuum robots to measure the environment contact forces directly. Haraguchi et al.[21] developed a force-sensing algorithm based on pneumatic cylinders and achieved three-axis sensing of translational forces. By gauging changes in magnetic fields, Hall Effect sensors have provided a force sensing method by measuring robot bending and deformation information [22]. However, all

This work was carried out while studying for an M.Sc. in Robotics and Computation at UCL and was supported by the Royal Society research grant RGS\$R1\231472

[1] Department of Mechanical Engineering, University College London, UK [2]Department of Computer Science, University College London, UK.

Source code and data for this paper can be found in : https://github.com/chendaxia777/Vision-force-estimation

of these methods necessitate auxiliary sensing apparatuses, potentially impinging on the robot's inherent flexibility and size—a drawback in applications demanding stringent size and flexibility constraints, like Minimally Invasive Surgery [7].

Emerging deep learning techniques have also prompted researchers to explore hybrid force estimation strategies that incorporate machine vision [23]. Haouchine et al.[24] present methods for vision-based force estimation in surgical robotic systems by detecting tissue deformation. However, this approach relies on external information from the environment, which is not easily accessible. Feng et al.[23] developed a method to estimate a robot's tip force by measuring tension and tendons' position. The contact force estimator, based on a recurrent neural network, is capable of estimating the tip force of a tendon-driven robot but is not suitable for use with pneumatic actuation or any systems employing low-level controllers.

In this study, we present a new and innovative tip force estimator derived from visual data and machine learning techniques. This estimator incorporates simple, arbitrarily placed cameras to capture the intricate deformation patterns the continuum robot exhibits. No additional sensors or calibration is required for the method, making it easy to be deployed. The core of our estimator relies on the strategic use of images gathered from the camera, combined with the known actuation pressure values of the robot. This combination serves as the input to a deep neural network, which subsequently computes the 2D external force exerted on the robot's tip.

The subsequent sections of this paper are organized as follows: In Section II, we delve into the foundational theories that underpin our force-sensing approach, shedding light on the fundamental principles that guide our methodology. Section III presents the experimental platform, data acquisition protocol, and deep neural network architecture we employ. Section IV showcases experimental outcomes, showcasing our method's advantages and bottlenecks. Finally, Section V encapsulates our findings and charts the trajectory for future research.

II. Theory

A general soft robot can be parametrized through the following three spaces [2]:

1) Joint Space ($\mathcal{Q} \in \mathbb{R}^{n_q}$): This encompasses the entire set of feasible actuator states (pressure in the context of this paper) within the defined design constraints of the robot.
2) Configuration Space ($\mathcal{C} \in \mathbb{R}^{n_c}$): The set of all attainable robot shapes.
3) Task Space ($\mathcal{T} \in \mathbb{R}^6$): The set of reachable end-effector pose, encompassing both positions and orientations.

Here n_q is the number of actuation variables (6, in this study), and n_c denotes the number of configuration parameters that full defines the shape of the soft robot (assumed to be unknown). These spaces are parameterized using their respective vectors: the joint space vector $\mathbf{q} = \begin{bmatrix} q_1, q_2, \dots \end{bmatrix}^\top \in \mathcal{Q}$, the configuration vector $\boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\psi}_1^\top, \boldsymbol{\psi}_2^\top, \dots \end{bmatrix}^\top \in \mathcal{C}$, and the end-effector pose vector $\mathbf{x}_e = \begin{bmatrix} x, y, \dots \end{bmatrix}^\top \in \mathcal{T}$.

The mapping from joint space to configuration space can be represented as $\boldsymbol{\psi} = F(\mathbf{q})$, while the mapping from configuration space to task space is represented as $\mathbf{x}_e = G(\boldsymbol{\psi})$. Consequently, the instantaneous direct kinematics from the configuration space to the twist space (derivative of end-effector pose) can be expressed as follows:

$$\dot{\mathbf{x}}_e = \mathbf{J}_{\mathbf{x}\boldsymbol{\psi}} \dot{\boldsymbol{\psi}} \tag{1}$$

Similarly, the instantaneous inverse kinematics from the configuration space to the joint space can be represented as:

$$\dot{\mathbf{q}} = \mathbf{J}_{\mathbf{q}\boldsymbol{\psi}} \dot{\boldsymbol{\psi}} \tag{2}$$

where $\mathbf{J}_{\mathbf{x}\boldsymbol{\psi}}$ and $\mathbf{J}_{\mathbf{q}\boldsymbol{\psi}}$ denote the mapping Jacobians from configuration space to twist space and joint space, respectively [25].

For the energy conservation to remain valid, the following relationship must hold statically [11]:

$$\mathbf{J}_{\mathbf{q}\boldsymbol{\psi}}^T \boldsymbol{\tau} + \mathbf{J}_{\mathbf{x}\boldsymbol{\psi}}^T \mathbf{W}_e = \nabla \mathbf{E} \tag{3}$$

where $\boldsymbol{\tau}$ denotes the actuation forces applied on the robot. $\mathbf{W}_e = \begin{bmatrix} \mathbf{F}_t^T & \mathbf{m}_t^T \end{bmatrix}^T$ represents the external wrench exerted on the tip of the robot, where $\mathbf{F}_t$ indicates the force and $\mathbf{m}_e$, the moment. $\nabla \mathbf{E}$ is the gradient of the elastic energy of the robot, which can be calculated when the stiffness $\mathbf{K}$ and deformation of the robot $\mathbf{S}_d \in \mathbb{R}^{n_c}$ are known [11].

Given the robot deformation $\mathbf{S}_d$ and input actuator pressure $\boldsymbol{\tau}$, the tip force applied on the robot can be fully determined:

$$\widehat{\mathbf{F}}_t = h(\mathbf{S}_d, \boldsymbol{\tau}) \tag{4}$$

where $\widehat{\mathbf{F}}_t$ stands for the estimated tip force, as illustrated in Figure 1. In this work, we assume that information about the robot deformation can be obtained from visual images and the function h can be approximated using a Convolutional Neural Network.

III. Methodology

A. Experimental Setup

The force prediction method introduced in this paper is implemented and validated using the STIFF-FLOP robot. The design and fabrication details of the robot is presented in detail in previous works [26], [27], [28]. The experimental setup features a two-segment STIFF-FLOP robot. Each robotic segment has a diameter of 25 mm and an overall length of 46 mm. To control the robot, an Arduino Due was used to regulate control voltages to six pressure regulators, generating pneumatic pressure (0–3 bar) to actuate the robot [29]. Two Lenovo cameras with a resolution of 1920×1080 were used to capture the deformation of the soft continuum robot, as shown in Figure 2(a). To meet the network's input requirements, the images captured from the cameras are pre-processed: they are resized to a resolution of 64×64 pixels and converted to grayscale as shown in

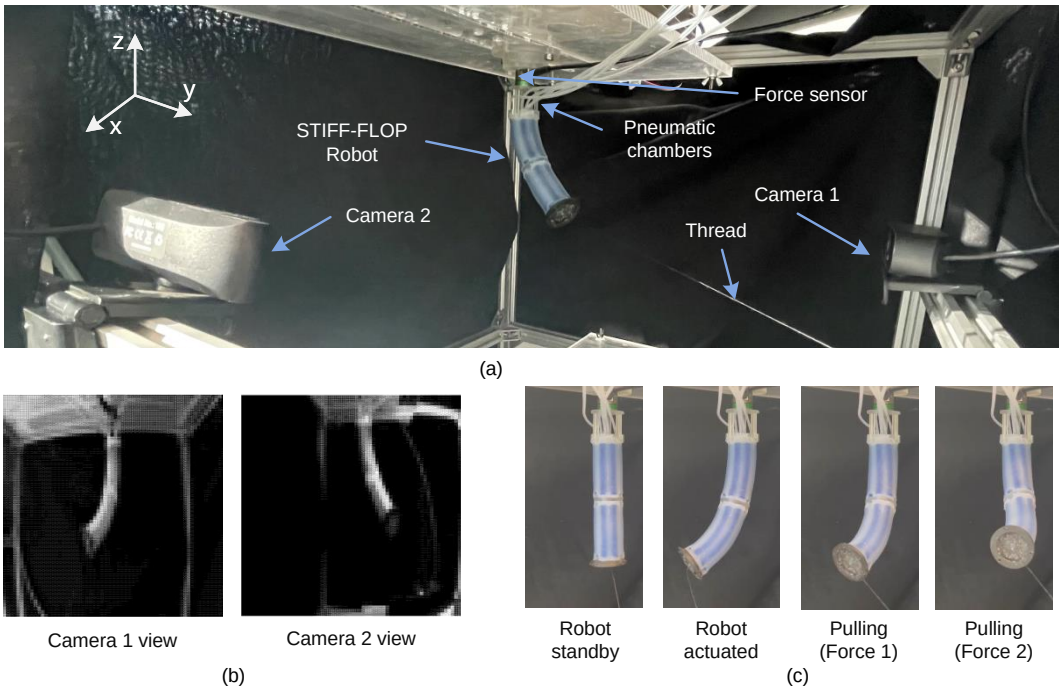

Fig. 2. Experimental platform and setup illustration. (a) Depiction of the experiment configuration, detailing the positioning of the STIFF-FLOP robot and the cameras. (b) Images resized to a resolution of 64×64 and converted to grayscale, matching the network input specifications. (c) Illustration of the robot's force application method through pulling.

Figure 2(b). A black thread was fixed to the robot's tip to apply forces. The robot would receive tip forces in different configurations by pulling the thread from various directions with various magnitudes. To obtain the ground truth force at the robot's tip without obstructing the cameras' field of view, a 6-axis force/torque sensor (IIT-FT17) was rigidly connected to the robot's base. This force sensor enables direct and reliable force measurements (error of ± 0.01 N) during the experiments. The control of the robot and the acquisition of measurements are done using MATLAB.

B. Data Collection

The pulling of the thread was done manually, and the overall procedure is illustrated in Figure 2(c). The experiment commenced with the robot being actuated by the application of a random pressure state. The force data over a duration of the first 10 seconds was measured, and its mean value was subtracted from the sensor data to compensate for gravitational effects.

After this process, the thread was manually pulled in different directions with varying force magnitudes. During this procedure, the cameras, and the force sensor continuously recorded data, capturing the robot's deformation and the exerted forces as ground truth. The data was sampled at 2 Hz. Each pressure setting was subjected to approximately a 60-second procedure with varying external forces applied. The training dataset in this study comprised of 60 such sets of deformation and force data, yielding a total of 6,415 data points. A separate validation set is also obtained with a different pressure profile for further tests.

C. Learning Architecture

To learn the mapping from deformation and pressure input to external forces, a deep neural network was used, as shown in Figure 3. The inputs to the network are greyscale images and a supersized actuation pressure matrix, both of size 64×64. The number of input images varied depending on the number of cameras involved. For the model utilizing information from both cameras, the input to the network was a 64x64x3 matrix comprising two images and the pressure matrix and the model with a single camera, had an input size of 64x64x2.

The proposed network architecture comprises five convolutional layers and three transposed convolutional layers. Each of these layers is followed by a Rectified Linear Unit (ReLU) activation function. To prevent overfitting, a dropout layer is included before the fully-connected layer with a drop out rate of 0.5. For increased robustness in performance, the original training dataset was augmented by introducing Gaussian noise. During each iteration of training, Gaussian noise, having a mean of 0 and a variance of 0.01, was added to the training image. This was done to improve the robustness of the network in real-world scenarios. The regression output layer produces a three-dimensional force output; however, our primary interest lies in the first two forces acting on the XY plane. This network was constructed and trained using the MATLAB Deep Learning toolbox. Training sessions utilized a mini-batch size of 516 and employed the Adam optimizer. The entire training process, conducted on a GeForce RTX 3060 GPU, took approximately 12 minutes.

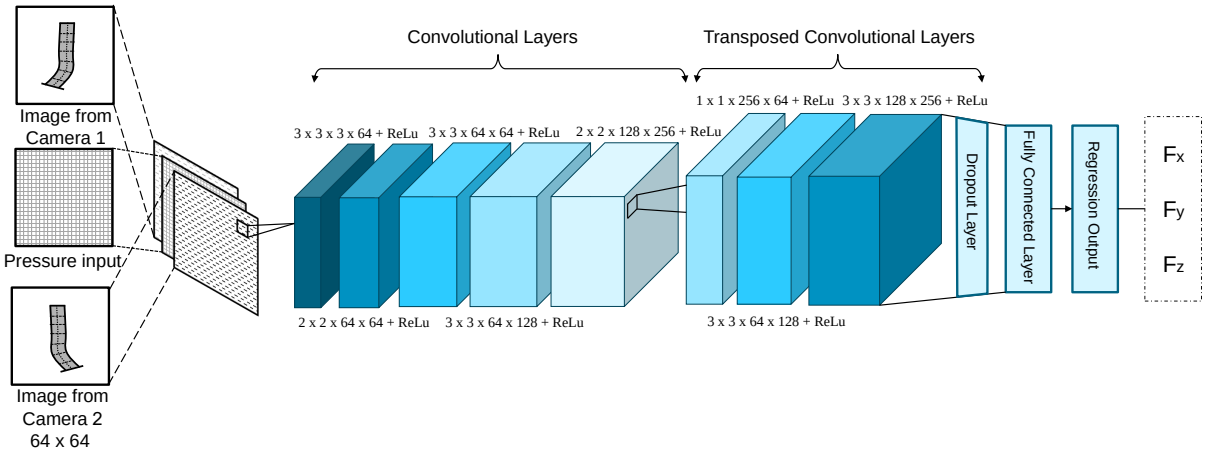

Fig. 3. The deep neural network architecture used in this work. The inputs to the network are images and a supersized actuation pressure matrix, both of size 64×64. The network outputs the force applied at the tip.

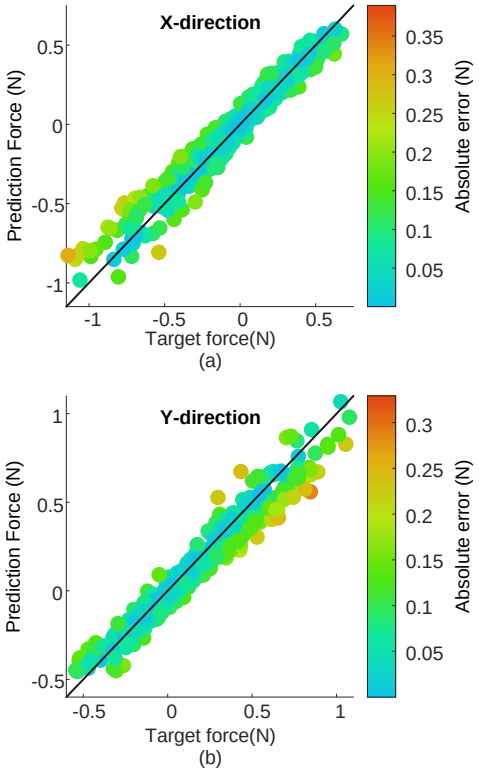

Fig. 4. Predicted force vs true force on the test set in the (a) X direction (b) Y direction. The data point colour represents the prediction error.

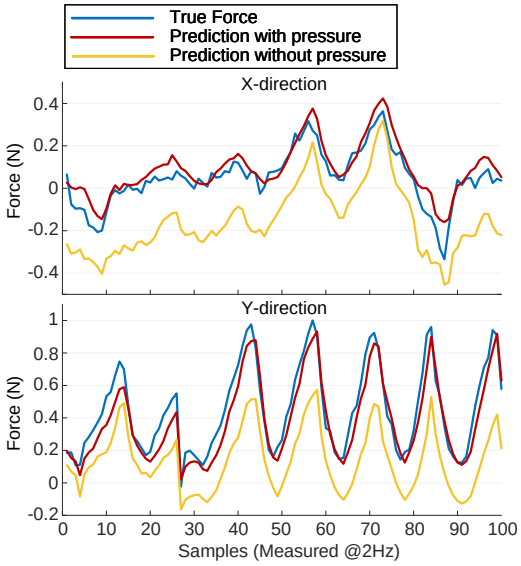

Fig. 5. Comparison of network predictions, with and without pressure information.

IV. Results

In this section, we present the performance of our vision-based force estimation model and evaluate its robustness to varying conditions. All the results shown are for the 2 camera system, unless specified otherwise.

A. Validation Results

To evaluate the accuracy of our model, the dataset was partitioned into a training set (90% of the data) and a testing subset (the remaining 10%). Figure 4 presents the network's prediction accuracy on the test data. It can be observed that the network exhibits the ability to estimate forces in both directions equally well, however, the performance deteriorates at extrema. We hypothesize that this is because of low representation in the training data, which can be rectified with more data at higher limits. The prediction results in the Z-direction is not presented here due to much smaller variations.

B. Analysis on Network Architecture

To analyse the relevance of pressure information for force prediction, we perform an experiment where the pressure matrix was removed from the network input. The subsequent network was trained using only the image data. The prediction performance is shown in Figure 5, revealing that the model without pressure information can still roughly estimate

TABLE I

PERFORMANCE IN PREDICTING FORCES APPLIED AT EITHER THE TIP OR THE BODY.

Mean absolute error and standard deviation of the predicted force (N)		
Test	Error on X direction	Error on Y direction
Tip Force (1-Camera)	0.0986 (0.11)	0.1911 (0.17)
Tip Force (2-Camera)	0.0593 (0.06)	0.1423 (0.14)
Body Force (2-Camera)	0.1979 (0.22)	0.2611 (0.31)

force patterns, but displays constant biases. When deprived of pressure information, the model must estimate the robot's initial actuation state purely from image data. Although it can obtain some information about the initial deformation state, it cannot capture it completely, leading to this observed bias.

To analyse the observability of the camera system, the network was trained using only one camera's images (camera-1 in Figure 2). Figure 6 shows the performance difference between using both cameras and a single camera. As expected, the performance in the Y direction is better when using images from both cameras, whereas the performance in the X direction is not significantly affected. This improvement can be attributed to the fact that camera-1 is in the Y direction, making it difficult to fully observe deformations in the Y direction. Hence, full observability of the system state is not necessary for force estimation and as the sensor quality improves the prediction accuracy increases. The presented model using visible spectrum images for force predictions, but these can also be replaced with other sensor sources that provide similar information. These can be from strain sensor data, X-ray imaging or magnetic sensors [30], [31].

C. Model Robustness

Thus far, the force estimation model has been evaluated under controlled conditions that closely resemble the training environment. Here, we evaluate the robustness of the predictions in varying conditions. As previously described, the training data was collected with force applied to the tip of the robot using a fixed thread. For our first study, we apply random forces to various locations on the robot using

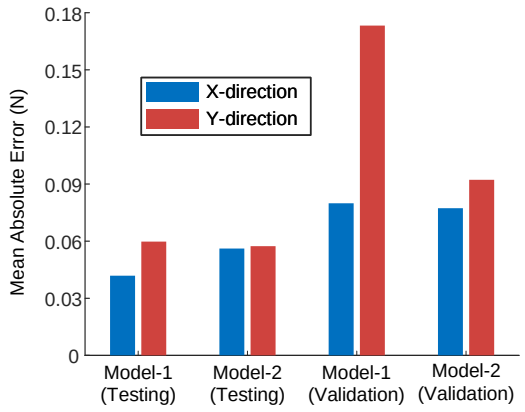

Fig. 6. Performance assessment of models based on camera input sources. Model-1 utilizes images only from camera-1 (See Figure 2), while Model-2 uses images from both cameras.

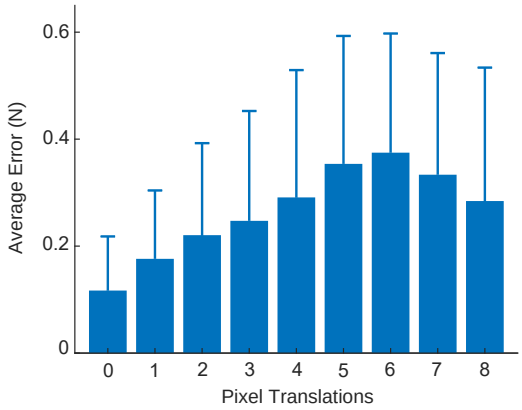

Fig. 7. Figure illustrating the model's performance as the image input is shifted. Images were translated by a certain number of pixels in both the X and Y directions.

TABLE II

FORCE PREDICTION ACCURACY ON DEFORMABLE OBJECTS

Mean absolute error and standard deviation (N)		
Material	X direction	Y direction
Rod	0.0593 (0.06)	0.1423 (0.14)
Sponge	0.0949 (0.12)	0.1137 (0.15)
Silicone rubber 1	0.0889 (0.10)	0.0887 (0.10)
Silicone rubber 2	0.1629 (0.20)	0.0956 (0.12)

a metal rod. The prediction force's mean absolute error and standard deviation is presented in Table I, comparing the performance of predicting forces applied at the tip versus forces applied elsewhere on the robot. The results reveal that predictions are reliable even when the visual scene is varied. However, the deformation modes obtained for point loads at different locations are not the same, leading to higher prediction errors. This could be reduced by collecting more varied force application data or by predicting moments rather than forces, which are distance independent.

Next, we study the effects of camera movement, a likely scenario in real-world applications. This is tested by virtually translating the images in both the X and Y directions. The resulting performance and variance of the prediction error are shown in Figure 7. The results show a sharp increase in the prediction error as the image is translated, with a slight drop observed after the images were moved more than 6 pixels. However, the variance of the error continued to increase, signifying that camera movement can significantly impact the accuracy of the network. Adding max pooling layers to the CNN architecture, further data augmentation (include random translations of the input data), and/or more advanced architectures like Spatial Transformer Networks (STNs) can be used to increase shift invariance [32].

Finally, we evaluate the force prediction accuracy in non-rigid surfaces to simulate real-world conditions. For comparison purposes, we use a metal rod and materials with different stiffness levels (See Figure 8). The external forces were applied in random directions and patterns, while ensuring that the force remained below the 0.5 N range.

The average prediction performance is shown in Table II

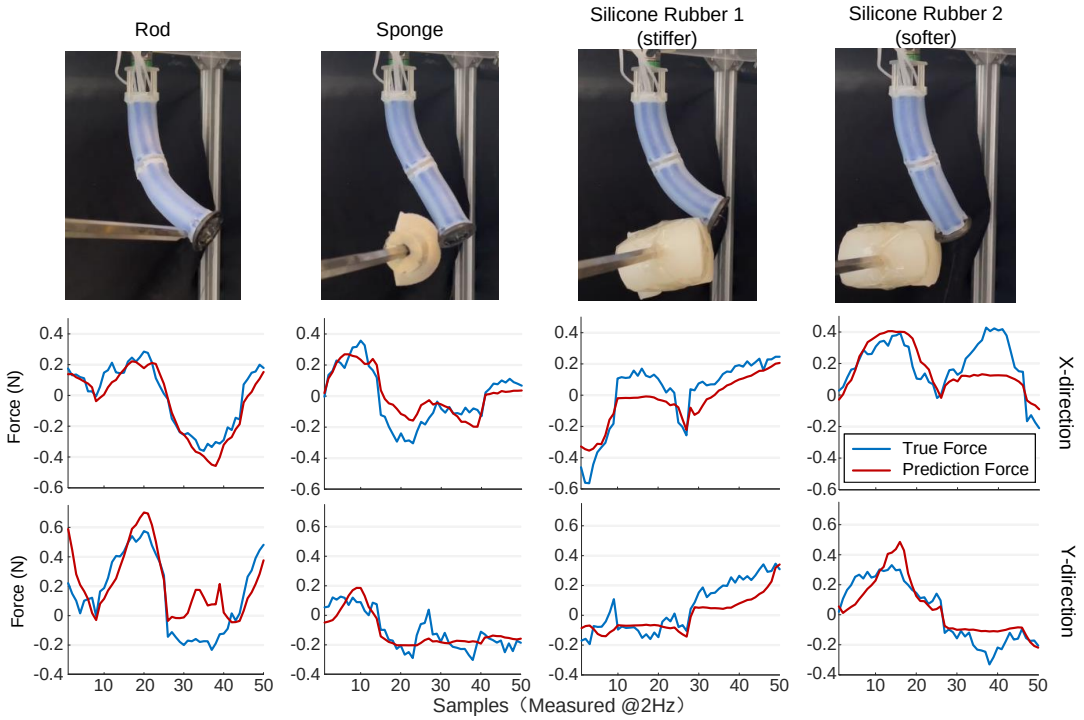

Fig. 8. Presentation of the force application setup, highlighting materials that could obstruct camera views and impact prediction performance. Forces were applied using various materials, including a metal rod, sponge, and silicone rubbers with varied stiffness. The top row illustrates force estimations in the X direction, while the bottom row shows performance in the Y direction.

and an illustration of the prediction model is shown in the lower part of Figure 8. The results indicate that the accuracy of force prediction declines as the external contact becomes softer. This decrease in accuracy may stem from two potential factors: first, the network's lack of training on visually novel objects; second, the broader force distribution from the soft object, leading to varied deformation in the soft robot.

V. CONCLUSION

This study presents a vision-based tip force estimation technique for soft continuum robots using deep learning that provides high accuracy without compromising the robot's structure. We tested this method on a two-segment STIFF-FLOP robot, utilizing two camera images to capture the robot's deformation. Our findings indicate that for optimal force estimation accuracy, the robot's actuation pressure is essential. We also enhanced the network's capability to function effectively even in the presence of obstacles within the images by adding visual noise while training, and validated the results using real-world experiments. Our vision-based force estimation model demonstrated a sensing precision of 0.05 N in the XY plane during testing, with a data collection and training time of only 70 minutes. Longer and varied training data should improve model performances.

While our experiments were conducted solely on the STIFF-FLOP robot, in theory, this method applies to various soft robots that exhibit deformation under external force. Investigations on other types of soft robots may be undertaken in future studies. Currently, the technique can only estimate the force at the robot's tip. Estimating forces along the robot's length is also theoretically possible but requires more experimental studies to be validated. Using other sources of deformation information like X-ray imaging or magnetic sensing could also be a possible future direction to be investigated.

REFERENCES

[1] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," *Nature*, vol. 521, no. 7553, pp. 467–475, 2015.

[2] P. E. Dupont, N. Simaan, H. Choset, and C. Rucker, "Continuum robots for medical interventions," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 847–870, 2022.

[3] N. Simaan, R. M. Yasin, and L. Wang, "Medical technologies and challenges of robot-assisted minimally invasive intervention and diagnostics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 465–490, 2018.

[4] M. Russo, S. M. H. Sadati, X. Dong, A. Mohammad, I. D. Walker, C. Bergeles, K. Xu, and D. A. Axinte, "Continuum robots: An overview," *Advanced Intelligent Systems*, vol. 5, no. 5, p. 2200367, 2023.

[5] J. Burgner-Kahrs, D. C. Rucker, and H. Choset, "Continuum robots for medical applications: A survey," *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1261–1280, 2015.

[6] H. Gao, X. Ai, Z. Sun, W. Chen, and A. Gao, "Progress in force-sensing techniques for surgical robots," *Journal of Shanghai Jiaotong University (Science)*, pp. 1–12, 2023.

[7] C. Shi, X. Luo, P. Qi, T. Li, S. Song, Z. Najdovski, T. Fukuda, and H. Ren, "Shape sensing techniques for continuum robots in minimally invasive surgery: A survey," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1665–1678, 2017.

[8] D. B. Camarillo, C. F. Milne, C. R. Carlson, M. R. Zinn, and J. K. Salisbury, "Mechanics modeling of tendon-driven continuum manipulators," *IEEE transactions on robotics*, vol. 24, no. 6, pp. 1262–1273, 2008.

[9] A. Gao, R. J. Murphy, H. Liu, I. I. Iordachita, and M. Armand, "Mechanical model of dexterous continuum manipulators with compliant joints and tendon/external force interactions," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 465–475, 2016.

[10] R. J. Webster III and B. A. Jones, "Design and kinematic modeling of constant curvature continuum robots: A review," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1661–1683, 2010.

[11] K. Xu and N. Simaan, "An investigation of the intrinsic force sensing capabilities of continuum robots," *IEEE Transactions on Robotics*, vol. 24, no. 3, pp. 576–587, 2008.

[12] D. C. Rucker and R. J. Webster, "Deflection-based force sensing for continuum robots: A probabilistic approach," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 3764–3769.

[13] M. Khoshnam, A. C. Skanes, and R. V. Patel, "Modeling and estimation of tip contact force for steerable ablation catheters," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 5, pp. 1404–1415, 2015.

[14] S. Hasanzadeh and F. Janabi-Sharifi, "Model-based force estimation for intracardiac catheters," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 1, pp. 154–162, 2016.

[15] A. Bajo and N. Simaan, "Hybrid motion/force control of multi-backbone continuum robots," *The International journal of robotics research*, vol. 35, no. 4, pp. 422–434, 2016.

[16] F. Khan, R. J. Roesthuis, and S. Misra, "Force sensing in continuum manipulators using fiber bragg grating sensors," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 2531–2536.

[17] H. Bai, S. Li, J. Barreiros, Y. Tu, C. R. Pollock, and R. F. Shepherd, "Stretchable distributed fiber-optic sensors," *Science*, vol. 370, no. 6518, pp. 848–852, 2020.

[18] K. R. Henken, J. Dankelman, J. J. van den Dobbelsteen, L. K. Cheng, and M. S. van der Heiden, "Error analysis of fbg-based shape sensors for medical needle tracking," *IEEE/ASME Transactions on mechatronics*, vol. 19, no. 5, pp. 1523–1531, 2013.

[19] Y. Noh, J. Bimbo, S. Sareh, H. Wurdemann, J. Fraś, D. S. Chathuranga, H. Liu, J. Housden, K. Althoefer, and K. Rhode, "Multi-axis force/torque sensor based on simply-supported beam and optoelectronics," *Sensors*, vol. 16, no. 11, p. 1936, 2016.

[20] R. L. Truby, C. Della Santina, and D. Rus, "Distributed proprioception of 3d configuration in soft, sensorized robots via deep learning," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3299–3306, 2020.

[21] D. Haraguchi, T. Kanno, K. Tadano, and K. Kawashima, "A pneumatically driven surgical manipulator with a flexible distal joint capable of force sensing," *IEEE/ASME Transactions on Mechatronics*, vol. 20, no. 6, pp. 2950–2961, 2015.

[22] M. Luo, Y. Pan, E. H. Skorina, W. Tao, F. Chen, S. Ozel, and C. D. Onal, "Slithering towards autonomy: a self-contained soft robotic snake platform with integrated curvature sensing," *Bioinspiration & biomimetics*, vol. 10, no. 5, p. 055001, 2015.

[23] 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, p. 17482, 2021.

[24] N. Haouchine, W. Kuang, S. Cotin, and M. Yip, "Vision-based force feedback estimation for robot-assisted surgery using instrument-constrained biomechanical three-dimensional maps," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2160–2165, 2018.

[25] K. Xu and N. Simaan, "Intrinsic wrench estimation and its performance index for multisegment continuum robots," *IEEE Transactions on Robotics*, vol. 26, no. 3, pp. 555–561, 2010.

[26] J. Fraś, J. Czarnowski, M. Maciaś, J. Główka, M. Cianchetti, and A. Menciassi, "New stiff-flop module construction idea for improved actuation and sensing," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 2901–2906.

[27] H. Abidi, G. Gerboni, M. Brancadoro, J. Fras, A. Diodato, M. Cianchetti, H. Wurdemann, K. Althoefer, and A. Menciassi, "Highly dexterous 2-module soft robot for intra-organ navigation in minimally invasive surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 14, no. 1, p. e1875, 2018.

[28] J. Shi, W. Gaozhang, and H. A. Wurdemann, "Design and characterisation of cross-sectional geometries for soft robotic manipulators with fibre-reinforced chambers," in *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*. IEEE, 2022, pp. 125–131.

[29] J. Shi, W. Gaozhang, H. Jin, G. Shi, and H. A. Wurdemann, "Characterisation and control platform for pneumatically driven soft robots: Design and applications," in *2023 IEEE International Conference on Soft Robotics (RoboSoft)*, 2023, pp. 1–8.

[30] E. J. Lobaton, J. Fu, L. G. Torres, and R. Alterovitz, "Continuous shape estimation of continuum robots using x-ray images," in *2013 IEEE international conference on robotics and automation*. IEEE, 2013, pp. 725–732.

[31] P. E. Dupont, B. J. Nelson, M. Goldfarb, B. Hannaford, A. Menciassi, M. K. O'Malley, N. Simaan, P. Valdastri, and G.-Z. Yang, "A decade retrospective of medical robotics research from 2010 to 2020," *Science robotics*, vol. 6, no. 60, p. eabi8017, 2021.

[32] M. Jaderberg, K. Simonyan, A. Zisserman, *et al.*, "Spatial transformer networks," *Advances in neural information processing systems*, vol. 28, 2015.