CathSim: An Open-source Simulator for Endovascular Intervention

Tudor Jianu[1], Baoru Huang[2], Minh Nhat Vu[3], Mohamed E. M. K. Abdelaziz[2], Sebastiano Fichera[4],
Chun-Yi Lee[5], Pierre Berthet-Rayne[6], Ferdinando Rodriguez y Baena[7], Anh Nguyen[1]

Abstract— Autonomous robots in endovascular operations have the potential to navigate circulatory systems safely and reliably while decreasing the susceptibility to human errors. However, there are numerous challenges involved with the process of training such robots, such as long training duration due to sample inefficiency of machine learning algorithms and safety issues arising from the interaction between the catheter and the endovascular phantom. Physics simulators have been used in the context of endovascular procedures, but they are typically employed for staff training and generally do not conform to the autonomous cannulation goal. Furthermore, most current simulators are closed-source, which hinders the collaborative development of safe and reliable autonomous systems. In this work, we introduce CathSim, an open-source simulation environment that accelerates the development of machine learning algorithms for autonomous endovascular navigation. We first simulate the high-fidelity catheter and aorta with the state-of-the-art endovascular robot. We then provide the capability of real-time force sensing between the catheter and the aorta in the simulation environment. We validate our simulator by conducting two different catheterization tasks within two primary arteries using two popular reinforcement learning algorithms, Proximal Policy Optimization (PPO) and Soft Actor-Critic (SAC). The experimental results show that using our open-source simulator, we can successfully train the reinforcement learning agents to perform different autonomous cannulation tasks. Our CathSim simulator is publicly available at https://github.com/tudorjnu/cathsim.

I. INTRODUCTION

ENDOVASCULAR intervention has been continuously evolving since the traditional approach of direct, open-cut surgery. It involves the use of a small incision that allows surgical equipment (such as catheters and guidewires) to be manoeuvred within the vasculature. This type of minimally invasive surgery (MIS) provides numerous advantages where the patient benefits from reduced blood loss, shorter recovery time, lower postoperative pain, and diminished inflammatory response compared to the traditional approaches [1]. In typical clinical conditions, the catheter and guidewire are navigated to the diagnosis zone through the use of fluoroscopy, a medical visualization procedure that obtains real-time X-Ray images from the operating theatre. Despite the relative advantages, endovascular intervention still presents some drawbacks such as lack of sensory feedback, surgeon

[1]Department of Computer Science, University of Liverpool, UK
[2]The Hamlyn Centre for Robotic Surgery, Imperial College London, UK
[3]Automation & Control Institute (ACIN), TU Wien, Austria
[4]Deparment of Mechanical, Materials & Aerospace Engineering, University of Liverpool, UK
[5]Department of Computer Science, National Tsing Hua University, Taiwan
[6]Honorary Fellow, University of Liverpool
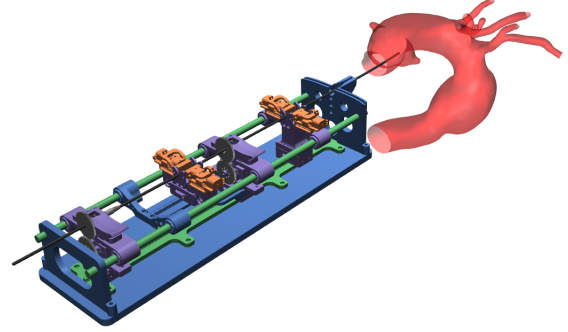[7]Mechatronicsin Medicine Lab, Imperial College London, UK.

Fig. 1. An overview of our **CathSim** - an open-sourced simulator for autonomous cannulation. The figure shows the follower robot navigating the catheter through an aortic arch.

exposure to radiation, and the need for highly dexterous manipulation [2].

To reduce the continuous risk of radiation imposed on the surgeon throughout the fluoroscopic procedure, many robotic systems with leader-follower teleoperation architecture have been proposed [3]–[5]. The surgeon actuates the leader device, from which, the information is mapped to the follower robot that executes the related action. The use of leader-follower robots allows the surgeon to perform the procedure remotely from a safe, radiation-free zone. Recent work has further focused on creating Magnetic Resonance (MR) safe robotic platforms [3], which eliminates the ionizing radiation exposure whilst allowing the soft tissue, such as the vasculature, to be visualized [6], [7]. Furthermore, in academic settings, the robotic system is developed to provide additional information to the surgeon through assistive features such as force/torque information [8], haptic feedback [9], and real-time segmentation and tracking [10]. Nonetheless, the procedure continues to be manually conducted by the surgeon due to the inherent technical, regulatory, ethical, and legal challenges [11].

Whilst recent robotic platforms for endovascular intervention demonstrate their assistive potential in the successful completion of the procedure, they share two problems: *i)* the lack of autonomy [21] and *ii)* increased duration of robotic procedure compared to its non-robotic counterpart [22]. Firstly, the surgeon operates within the tridimensional space of the vasculature whilst relying on the information provided by two-dimensional fluoroscopic images and haptic feedback. Additionally, the surgeon has to avoid inflicting

TABLE I
ENDOVASCULAR SIMULATORS COMPARISON

Simulator	Physics Engine	Catheter	Guidewire	Force Sensing	Open-source
Molinero *et al.* [9]	Unity Physics [12]	Discretized	✗	Vision-Based	✗
Karstensen *et al.* [13]	SOFA [14]	Timoshenko Beam theory [15]	✗	✗	✗
Behr *et al.* [16]	SOFA [14]	Timoshenko Beam theory [15]	✗	✗	✗
Omisore *et al.* [17]	CopelliaSim [18]	Unknown	✗	✗	✗
Schegg *et al.* [19]	SOFA [14]	Timoshenko Beam theory [15]	✗	✗	✗
CathSim (ours)	MuJoCo [20]	Discretized	✗	✓	✓

Note: Whilst Timoshenko Beam catheter plugin is open-sourced, the whole simulation environment is not.

extensive damage to the vasculature system, thus operating under mentally strenuous conditions which can be diminished through the automation of the procedure. A more autonomous surgery would ideally inflict little damage whilst operating in a timely manner. However, this is not a trivial task in practice as it requires a complex vision, learning, and control system that can guarantee the safety of the procedure [21].

Recent developments in machine learning promise a greater degree of autonomy in many robotic systems. Such systems leverage deep learning architectures such as convolutional neural networks [23], [24], long-short-term memory [25]–[27], and generative adversarial imitation learning [10] to facilitate force estimation and catheter segmentation. Whilst those systems confer a lower level autonomy through the use of robotic assistive features, the higher level autonomy is left unaddressed. However, the navigation task has been addressed through many works [13], [17], [22]. The environment employed by those works makes tradeoffs between the use of physical and virtual environments, and they generally rely on closed-source environments which cannot be replicated by fellow researchers. Furthermore, albeit the use of simulated environments, they generally do not adhere to the Reinforcement Learning (RL) paradigm.

In this work, our goal is to provide a new minimally invasive surgery environment for endovascular procedures. We aim at facilitating the development of endovascular autonomy through the provision of a standardized environment that confers familiarity to the machine learning community. As such, we propose **CathSim**, a real-time simulation environment for autonomous cannulation based on MuJoCo [20]. We choose MuJoCo as the base simulator as it is a real-time and accurate physics engine that facilitates optimal control applications, hence well-suited for our task. Fig. 1 shows the overview of our simulator. We summarize our contributions and the potential usage of our simulator as follows:

1) We propose **CathSim**, a new open-source simulation environment for endovascular procedures.
2) We implement the baseline and provide the benchmark for autonomous cannulation tasks in our simulator using two popular RL algorithms.
3) With real-time force sensing and high-fidelity visualization, our open-source simulator can be used in various tasks such as medical training or Augmented Reality (AR) and Virtual Reality (VR) applications.

II. RELATED WORK

Simulation Environments. Research within the simulation of environments for minimally invasive surgery divides the simulation level into four distinct categories: synthetic, animal, virtual reality and human cadaver [28] where each possesses its own limitations and advantages [29]–[31]. They mostly concentrate on the skill development of the trainee [28], [32], or the development of assistive features such as haptic feedback [9], or rely on physical materials which do not offer an RL-compliant environment. Recent research has been undertaken using synthetic simulators such as a high fidelity phantom [22] by means of imitation learning, whereas other research used the SOFA [14] simulation environment [33] and tested on a bi-dimensional synthetic phantom. Whilst the literature is advancing, the physical or closed-source nature of such simulators hinders the ability for collaborative improvement.

Autonomous Cannulation. The advancement of machine learning allows preliminary results of autonomous cannulation. In [22], generative adversarial imitation learning [34] has been used to perform catheterization tasks. The authors in [17] used deep RL in a simulated environment to perform robotic axial navigation for coronary interventions. Cho *et al.* [35] created an image-based automatic control approach for controlling the guidewire. Kweon *et al.* [36] applied deep RL for guidewire navigation in a coronary artery phantom. Omisore *et al.* [17] used a sample-efficient deep RL with episodic policy transfer. Within a two-dimensional phantom, Karstensen *et al.* [13] deployed a simulation-trained RL agent into a bi-dimensional endovascular phantom. Whilst the recent results are promising, the bulk of research is situated on the lower level of autonomy scale [37] using closed-source platforms whilst the complete navigation throughout the vascular system is yet to be achieved.

In Table I, we show a detailed comparison of current endovascular simulators. **CathSim**, in contrast to other simulators, provides an open-source environment that is well-suited to for training autonomous agents using different machine-learning approaches. Based on MuJoCo's [20] framework, **CathSim** offers an advanced simulation environment for real-time applications. Furthermore, our simulator provides real-time force sensing capability and high-fidelity realistic visualization of the aorta, catheter, and endovascular robots. In practice, **CathSim** can be used to train RL agents or serve as a practice platform for healthcare professionals.

III. THE CathSim SIMULATOR

Our **CathSim** environment has three components: *i)* the follower robotic model for endovascular procedures [7], *ii)* the aortic arch phantoms, and *iii)* the catheter. Our simulator enables real-time simulation and supports the training of state-of-the-art learning algorithms.

A. Robot Simulation

In this work, we aim at transferring CathBot [3] to simulation with the purpose of autonomous agents training. We choose CathBot as it is a state-of-the-art robot and is not bounded by a commercial licence. The design of CathBot follows the popular leader-follower architecture with the leader robot uses haptic feedback generated by the catheter's interaction with the environment through the navigation system [9] and maintains an intuitive control that replicates human motion patterns such as insertion, retraction, and rotation. The follower robot mimics the leader motion, and it is made up of two pneumatic linear motors for translation, one pneumatic rotating stepper motor, and two pneumatic J-clamps for clamping the instrument while performing translational motions.

Given the linear mapping between the leader and the follower robot in CathBot's design [3], we focus only on simulating the follower robot for simplicity. We simulate the follower robot by constructing four modular platforms that are attached to the main rail. On two of those platforms, a pair of clamps is set to secure the guidewire in place during the translational movements, whilst on the other two, rotary catheter and guidewire platforms are attached for performing the angular motions. The parts that account for the translational movements on the main rail as well as the clamps are joined using prismatic joints. Furthermore, revolute joints are used to bind the wheels, thus providing the catheter with rotational movements. The rotational aspect of the catheter implies frictional movement. That is, when the rotational movement is actuated, the clams lock the catheter in place and rotate it. A similar procedure is carried out when the catheter is linearly displaced. However, this friction-reliant rotation is difficult to undertake in simulation, and therefore we assume a perfect motion throughout the system and actuate the joints directly.

We simulate the follower of the CathBot robot together with the aortic arch phantoms and the catheter. We chose to model these elements using MuJoCo's [20] physics engine, given its stability and computational speed while enabling real-time interactions.

B. Aorta Simulation

To simulate the aorta, we first scan the silicone-based, transparent, anthropomorphic phantom, of the Type-I and Type-II aortic arch model (Elastrat Sarl, Switzerland) to create high-fidelity 3D mesh models. The concave mesh is then decomposed into a set of nearly convex surfaces using the volumetric hierarchical approximate decomposition [38] resulting in $1,024$ convex hulls for Type-I aortic arch and 470 convex hulls for Type-II aortic arch. The difference
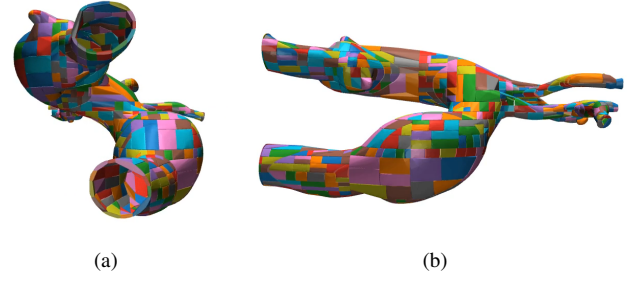

(a)   (b)

Fig. 2. The collision property of the aorta is inferred through the decomposition of the aortic arch into a series of convex hulls, whilst the visual properties are given by scanned high fidelity silicone-based, transparent, anthropomorphic phantoms.

in the number of convex hulls is given by the difference in the measure of concavity of the two meshes [39]. The convex hull is used to model the collision, as it aids the computational process and allows the use of soft contacts by the physics engine [20]. We show the aorta simulation in Fig. 2.
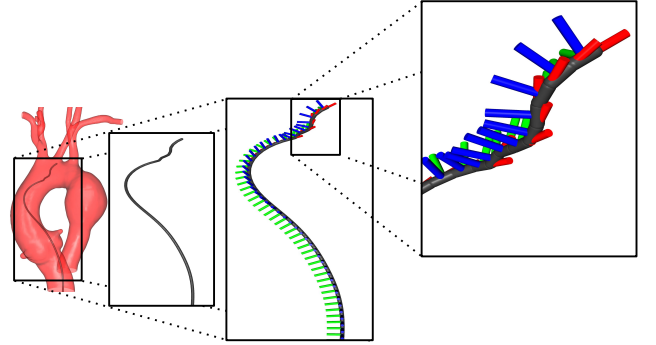
C. Catheter Simulation

Fig. 3. The catheter's continuous properties are approximated by a series of rigid bodies linked through revolute joints. The kinematic chain can be visualized through the frames of each individual body, where the last ten links are actuated by internal micro-motors.

One form of modelling the catheter represents the discretization of the continuous shape into a series of rigid bodies, interconnected by revolute or spherical joints [5]. This approach has been further categorized into discrete and serpentine methods, where the latter differs by having shorter rigid segments and, implicitly, a higher number of links [40]. The discretization approach has been proven to confer reasonable accuracy for shape prediction of the continuum robot [41]. In contrast to the discretization approach, the continuous curve estimation methods [5] assume that the described object is composed of an elastic backbone. The constant curvature model describes the continuum robot geometry with a finite number of mutually tangent curved segments, each with a constant curvature along its length or a variable curvature approach that takes into account the different curvatures across the backbone of the robot [5].

In this work, to enable real-time performance, we model the catheter based on a discretization approach by creating

a serpentine-like model in which its continuous shape and deformation are approximated by a series of rigid bodies and revolute joints. A prismatic joint was added for mimicking the direct linear movement of the catheter. Similarly to [42], we limit the actuation to the tip of the catheter, where the revolute joints are intrinsically actuated by motors. This approach allows the use of standardized modelling based on Newton-Euler equation [43] which is a good approximation of the inherent continuous shape that benefits from the computational efficiency of the method. The catheter model is shown in Fig. 3.

D. Contact Simulation

As our **CathSim** is based on MuJoCo [20], the contact between the aorta and the catheter are simulated using point contacts. In practice, point contact is defined geometrically as a point between two geometric objects and a spatial frame centred at that point in a global coordinate frame. The first axis of this frame is the contact normal direction, while the two other axes define the tangent plane. The contact distance is then used to determine if the penetration happens (i.e, the contact distance is positive if two geometric objects are separated, zero when they are in contact, and negative when they penetrate) [20].

IV. REINFORCEMENT LEARNING FOR AUTONOMOUS CANNULATION

We consider the task of autonomous cannulation, where our system represents an episodic Partially Observable Markov Decision Process (POMDP). Our agent, represented by the catheter, interacts with an environment $\mathcal{E}$ represented by an aortic arch type. At each time step t, the agent, receives an observation s_t, chooses an action $a_t(s_t)$, receives a reward $r_t(s_t, a_t)$ and arrives in a new state s_{t+1}. The episode terminates when the agent reaches the goal position within the aorta $g \in \mathcal{G}$.

A. Observations

We consider three types of observations for the experiments, namely Internal, Image, and Sequential.

Firstly, we employ Internal observation by assuming that the agent has access to as much information as possible. As such, we provide our system with the following information:

- position and velocity
- centre of mass inertia and velocity
- actuators generated force
- external forces generated on body

Secondly, we use the Image observation by positioning a virtual RGB camera of $128 \times 128 \times 3$ resolution on top of the aortic phantom. We map the image from RGB to greyscale as in [32], thus imitating the X-ray images and diminishing the dimensionality. This observation space simulates the clinical procedure, where the surgeon observes the fluoroscopic images. We could increase the image dimension to a higher resolution, although through our experiments, we observed that a higher image resolution does not increase

performance but increases the computational time and the memory requirements when using a replay buffer.

Thirdly, due to the difficulty of inferring the catheter actions given one image as in the "Image" observation setup, we consider the "Sequential" observation by inserting the temporal dimension through the concatenation of three subsequent images $\{s_{t-2}, s_{t-1}, s_t\}$. This observation would take into account the temporal domain of the catheter action and potentially provide more information to the RL agent.

B. Actions and Rewards

The actions are represented by a vector $a_t \in \mathbb{R}^{21}$, where the first 20 elements are associated with motors that actuate the revolute joints inside the tip of the catheter plus the prismatic joint responsible for the translational movement. Furthermore, the actions are normalized within an $[-1, 1]$ interval, giving a space of $a_t \in [-1, 1]^{21}$.

Given the sparsity of the reward function, we chose to convey more spatial information through reward shaping. Considering the navigation task, we provide the agent with an informational reward regarding the distance towards the target by computing the negative Euclidean distance between the head of the catheter h and the goal g, such as $d(h, g) = ||h - g||$. If the agent is within a distance δ of the goal, the episode terminates and the agent receives a further reward of $r = 10$. The distance of $\delta = 8$ mm was selected as within this distance, the catheter tip is fully inserted within the artery.

$$r(h_t, g) = \begin{cases} 10 & \text{if} \quad d(h, g) \leq \delta \\ -d(h, g) & \text{otherwise} \end{cases} \quad (1)$$

Whilst this reward function assists in agent convergence, it is also prone to local minima. An example would be the erroneous insertion of the catheter in another artery. In this case, the catheter would have to increase the distance from the target to achieve the goal objective.

C. Network Architectures

Considering the continuous action representation, we employ two state-of-the-art RL algorithms, namely PPO [44] and SAC [45] with the parameters utilized in [46]. A Multi-Layered Perceptron (MLP) based policy is used for the Internal observation and a Convolutional Neural Network (CNN) is used for the Image and Sequential observation. Note that, for simplicity, we follow [47] and use CNN for both Image and Sequential observations as they only have one and four input channels respectively where each channel represents a greyscale image. In our implementation, the MLP has two hidden layers of sizes 64 with a tanh activation function. The CNN has three convolutional layers with a ReLU activation function. The ADAM optimizer is used to train all networks with a learning rate of 0.0003.

V. EXPERIMENTS

In this section, we perform intensive experiments to validate our **CathSim**. We start with the experiment to verify

TABLE II

RESULTS SUMMARY

Aorta	Observation	Reward		Mean Force (N) ↓		Max Force (N) ↓		Success %	
		BCA	LCCA	BCA	LCCA	BCA	LCCA	BCA	LCCA
Type-I	PPO-Internal	-70 ± 28	-154 ± 66	0.011 ± 0.021	0.008 ± 0.015	0.121 ± 0.095	0.123 ± 0.089	53	07
	SAC-Internal	-54 ± 36	-158 ± 83	0.007 ± 0.014	0.006 ± 0.015	0.059 ± 0.047	0.095 ± 0.039	77	07
	PPO-Image	-65 ± 52	$\mathbf{-140} \pm 96$	$\mathbf{0.005} \pm 0.007$	0.007 ± 0.016	$\mathbf{0.048} \pm 0.020$	0.076 ± 0.059	83	**37**
	SAC-Image	-391 ± 157	-196 ± 5	0.024 ± 0.035	$\mathbf{0.003} \pm 0.003$	0.258 ± 0.094	$\mathbf{0.059} \pm 0.028$	00	00
	PPO-Sequential	$\mathbf{-57} \pm 36$	-336 ± 55	0.006 ± 0.009	0.010 ± 0.015	0.045 ± 0.022	0.094 ± 0.036	**97**	00
	SAC-Sequential	-200 ± 20	-227 ± 12	0.003 ± 0.003	0.003 ± 0.003	0.043 ± 0.009	0.061 ± 0.022	00	00
Type-II	PPO-Internal	-52 ± 48	-22 ± 30	0.016 ± 0.023	0.011 ± 0.015	0.118 ± 0.103	0.068 ± 0.040	50	87
	SAC-Internal	$\mathbf{0} \pm 0$	-67 ± 30	0.014 ± 0.010	0.010 ± 0.016	$\mathbf{0.030} \pm 0.007$	0.121 ± 0.044	**100**	27
	PPO-Image	-20 ± 47	$\mathbf{-8} \pm 10$	0.017 ± 0.025	0.015 ± 0.022	0.046 ± 0.047	0.066 ± 0.054	87	**97**
	SAC-Image	-68 ± 91	-54 ± 7	$\mathbf{0.004} \pm 0.005$	$\mathbf{0.005} \pm 0.005$	0.042 ± 0.013	$\mathbf{0.050} \pm 0.008$	73	03
	PPO-Sequential	-47 ± 49	-80 ± 92	0.011 ± 0.017	0.011 ± 0.018	0.068 ± 0.059	0.066 ± 0.050	57	67
	SAC-Sequential	-99 ± 18	-85 ± 15	0.005 ± 0.007	0.006 ± 0.008	0.055 ± 0.017	0.100 ± 0.023	03	00

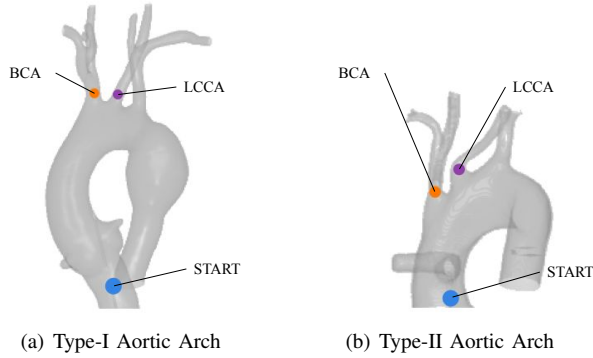(a) Type-I Aortic Arch   (b) Type-II Aortic Arch

Fig. 4. The starting configuration. The figure depicts the navigation task employed in the Type-I and Type-II Aortic Arches. The catheter is initially situated within the ascending aorta with the task of navigating towards the brachiocephalic artery (BCA) or the left common carotid artery (LCCA). The task finishes when the tip of the catheter is situated within a proximity of 8 mm of the targets.
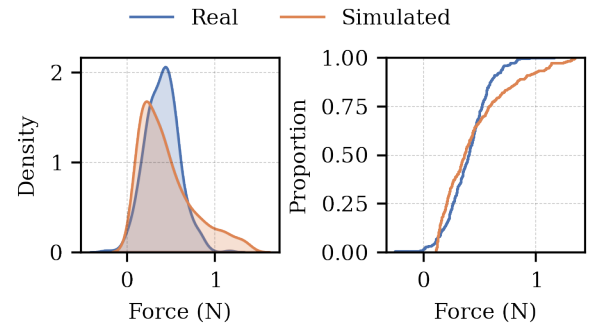

Fig. 5. The forces distributions are visualized through kernel density estimation (left) as well as through the cumulative distribution function (right). The figure depicts the similarities between the two distributions, where the empirical distribution gathered from the simulator shows a heavier tail than the normal distribution. Within the right subfigure, it can be observed that the cumulative distributions intersect at the 0.5 threshold, which signifies that the two distributions have a similar median ($\tilde{x}_1 \sim \tilde{x}_2$).

whether our simulator can mimic the behavior of the real-robot (CathBot). We then demonstrate how **CathSim** can be used for autonomous cannulation tasks using RL algorithms.

A. Simulator Validation

1) Setup: We assess the validity of our simulator, considering the distribution of forces generated throughout cannulation of the brachiocephalic artery (BCA). Ideally, we want to compare the force measured by our simulator and the force measured in the real-world experiment setup [3]. Note that, in the experiments conducted in [3], a load cell (Mini40, ATI Industrial Automation, Apex, NC, USA) was used to capture the force generated by the interaction of the instruments with the silicon phantom.

To extract the force from our simulator, we manually perform the cannulation using the keyboard, then in each simulation time step, we extract the collision points between the catheter and the aorta along with the tridimensional force expressed as the normal force f_z and frictional forces f_x and f_z. We further compute the magnitude of the force given the previous, such as given a time instance t, the magnitude of the force f_t is given by

$$f_t = \sqrt{f_{x,t}^2 + f_{y,t}^2 + f_{z,t}^2} \qquad (2)$$

We proceed by offering a comparison between the observed empirical distribution and a normal distribution derived from the real experiments conducted in [3]. We sample the same number of samples generated in our simulation and the real experiment in [3] for a fair comparison. We use the distributions to generate a cumulative distribution function $F_X(x) = P(X \le x)$, which can be visualized in Fig. 5. We further assess the normality of the sampled data, followed by a comparison of the two given distributions.

2) Results: We begin our comparison with the Shapiro-Wilk test of normality on the data extracted from the simulator and, given a p-value of $p \approx 7.195 \times 10^{-17}$ and a statistic of 0.878, we conclude that the sampled data does not represent a normal distribution $X(\mu, \sigma)$. This can be visualized in Fig. 5, where the sampled data is plotted against a normal distribution. Furthermore, we assess the homoscedasticity of the sample distribution and normal distribution by using Levene's tests, which results in a statistic of 40.818 and a p-value of $p \approx 2.898 \times 10^{-10}$, therefore, concluding that $\sigma_1^2 \neq \sigma_2^2$. Given the previous statistics (i.e.

non-normal distribution and unequal variances), we select the non-parametric Mann-Whitney test to compare the given distributions. The resulting statistic given the test is 76076, with a p-value of $p \approx 0.445$. Given that the p-value is higher than the threshold of $p = 0.05$, we can conclude that the differences in the distributions are merely given to chance and therefore the distributions can be considered as being part of the same population and thus convene that the force distribution of our simulator closely represents the distribution of forces encountered in the real-life system. Therefore, we can see that our **CathSim** successfully mimics the behavior of the real-world system.

B. Reinforcement Learning Results

1) Setups: We consider the autonomous catheterization of two principal arteries into the aortic arch, namely the brachycephalic artery (BCA) and the left common carotid artery (LCCA). Within both setups, we position the catheter tip at the starting locations within the ascending aorta and terminate the training when the catheter is fully inserted within the artery. We follow the same procedure for both the Type-I and the Type-II aortic arches (Fig. 4).

We trained the model on a series of tasks using the setup outlined in Section IV-B for a total of $600,000$ time steps. Each episode began with the catheter being randomly displaced by 1 mm. The agent then navigated through the ascending aorta, taking actions to reach the goal state, which was not specified. If the goal state was not reached after $2,000$ steps, the episode was terminated. At each step, we extracted information about the contact points between the aorta and the force exerted on it, as given by Eq. 2. This information was further used to create a 2D image overlaid on top of the RGB virtual camera image of the aorta, with each pixel representing the value of the force at that point for visualization purposes. The resulting heatmap is shown in Fig. 6. We conducted an evaluation of the model's performance over 30 episodes, collecting data on the force interactions and rewards for each episode. The average results are summarized in Table II, and the distribution of forces is shown in Fig. 7. The maximum force for a given episode with n samples is given by $\max_{t \in n} f_t$, and the mean force is calculated as

$$\overline{f} = \frac{\sum_{t=1}^{n} f_t}{n}$$

The training time varied from a minimum of 1 hour whilst training PPO to 8 hours whilst training SAC. Overall, PPO training is approximately five times faster than SAC training.

2) Quantitative Results: Using the extracted forces for each time step, we compute the mean force of the in-contact zones, as well as the maximum force exerted on the aorta during an episode. We then compute the average across the 30 episodes for both the maximum and the mean force. Furthermore, we compute the mean reward over the evaluated episodes and display the results in Table II. All observations perform similarly, showing close-by results in

(a) Type-I Aortic Arch - BCA   (b) Type-II Aortic Arch - BCA

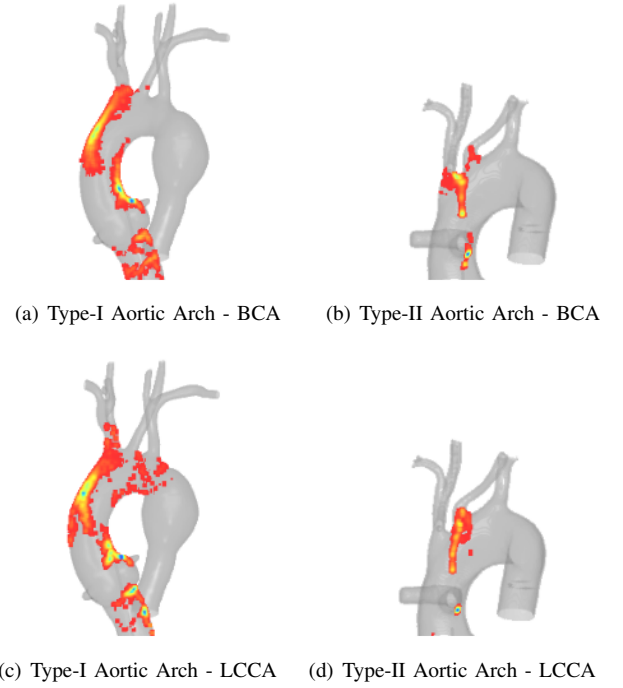(c) Type-I Aortic Arch - LCCA   (d) Type-II Aortic Arch - LCCA

Fig. 6. The force heatmap extracted from the interaction between the catheter and the aortas when undertaking the navigating towards the BCA target on the top row, and LCCA target on the second row. The catheterization procedure exerts a greater force on the first aorta where complicated manoeuvres are required to reach the targets, whilst on the second aorta it generates less force.

terms of force. Type-I Aortic Arch experiments show that PPO relying on a sequential observation space achieved the greatest reward when cannulating the BCA target (-57 ± 36), although it shows the least performance when the target is LCCA (-336 ± 55). The success rate mimics the reward, where the highest success rate resulted from the former (97%) and no success in the latter (0%). A more coherent reward has been achieved whilst using a singular image observation, where the cannulation of LCCA presents the greatest success (34%). In contrast, the cannulation of BCA presents close results to the sequential observation (83%). The catheterization of the Type-II aortic arch denotes an easier cannulation procedure, given the cannulation success rates. The task represents diminished wall reliability for undertaking the manoeuvre, thus representing a more direct catheterization procedure. This phenomenon can be also observed in Fig. 6, where the catheter exerted a greater amount of force on the aortic walls in order to reach the designated target. Overall, from our experiments, PPO shows better suitability for the task, especially when the observation is image-based.

Overall, the cannulation of Type-II aortic arch presents better results compared to Type-I. This is because the anatomy of Type-II aortic arch presents fewer complexities during the catheterization process. The Type-II aortic arch presents a shorter distance from the sino-tubular junction to the branching points of the head vessels. This anatomy provides a more direct path for catheterization, making it easier to access and navigate the catheter through the
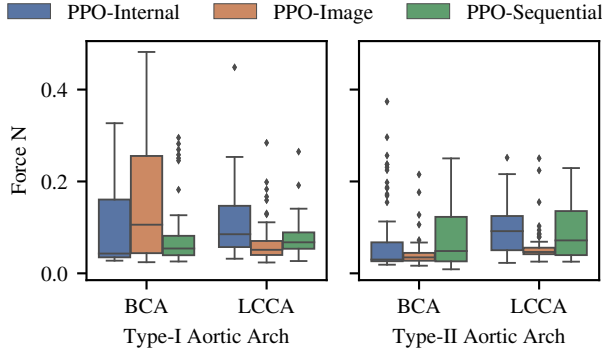
Fig. 7. The max forces generated by the catheter on the aorta within each of the evaluation episodes during the PPO training are shown. The left figure depicts the results gathered within the Type-I Aortic Arch where the catheter has greater interaction with the aorta, whereas the one on the right shows the results gathered within the Type-II Aortic Arch where the navigation is more direct.

ascending aorta. In contrast, Type-I aortic arch has a more complex anatomy, with a more arched ascending aorta. This can make it difficult for the catheter to reach the intended target vessel due to the longer and more tortuous pathway. Therefore, the anatomy of the Type-II aortic arch provides a more straightforward catheterization path, resulting in a higher success rate than the Type-I aortic arch.

3) Qualitative Result: We use the force frames extracted from each time step and compute the mean of the in-contact regions across all episodes. We further superimpose the heatmap over the phantom model using the projection matrix from the simulator, and we repeat for all the target configurations, resulting in the force heat maps visible in Fig. 6. The heatmap shows that the catheter applies a greater force at the points where bending actions are required to reach the target. For example, in Fig. 6 (c), the catheter bends on the ascending aorta to reach the brachiocephalic and left common carotid artery. Also, it is visible that the agent is prone to overshoot the LCCA target and deviate towards the descending aorta. Within the Type-II Aortic Arch, the catheter applies a greater force within the separating wall between BCA and LCCA, especially when the target is BCA as the catheter has to bend towards the target as it is pushed, contrary to LCCA where the forward motion with a slight descend. Overall, The Type-I Aortic Arch presents a more difficult environment than the Type-II Aortic Arch.

We further display the results of the force interaction in Fig. 7. The internal observation depicts the greatest amount of force and the greatest interquartile range in all cases except within the Type II aortic arch when performing the navigation towards the LCCA target. Additionally, it can be visualized that most of the points have a narrow interquartile range. However, outliers that exert a greater amount of force are present, with forces achieving higher than 0.4N.

At each time step, we compute the mean reward of the last 100 steps and display the results in Fig. 8. The algorithms show continuous learning with slight convergence. Within all
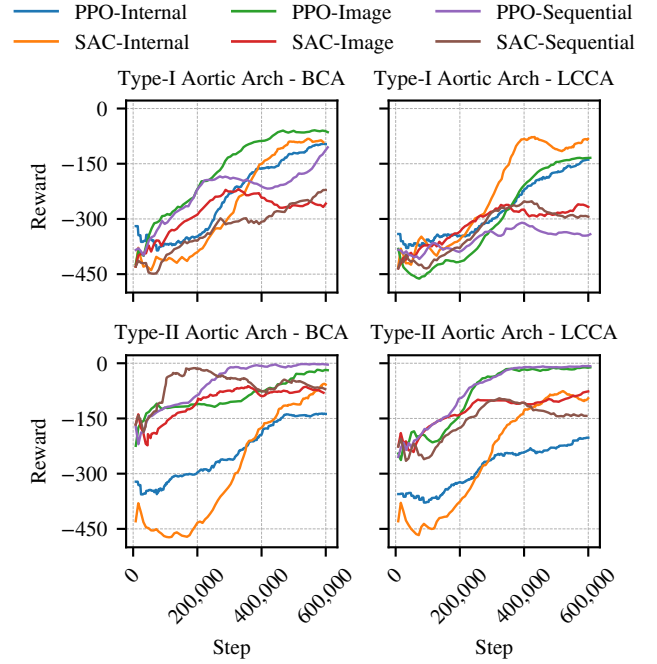

Fig. 8. The rewards when training autonomous agents using an on-policy algorithm (PPO) and an off-policy algorithm (SAC). In general, we observe that PPO performs better than SAC. The image observation consistently achieved good results when using PPO. Furthermore, the sequential observation shows the least amount of progress compared with the other observation spaces.

the environments, PPO with an image observation obtained the highest or close to the highest reward, thus showing that an on-policy algorithm, generally performs better for the given task. Furthermore, it shows that an image-based observation that correlates to the medical procedure undertaken in real clinical scenarios is capable of reaching the most reward. Whilst using the internal observation space, the off-policy algorithm (i.e. SAC) performed reasonably well, managing to obtain the highest reward within the Type-I aortic arch when cannulating the LCCA target whilst obtaining competitive results within the other tasks. Within all cases and coherent with the evaluation undertaken in the previous section, the sequential observation space yielded the least reward.

VI. CONCLUSIONS AND FUTURE WORK

We have presented **CathSim**, an open-source simulation environment that provides the community with a benchmarking platform for autonomous endovascular navigation. Using the **CathSim** simulator, we can develop and test different algorithms for autonomous cannulation without the need for physical robots. We believe that our work will bridge the gap between theory and practice, providing a development and benchmarking platform for computer scientists and roboticists, as well as a training platform for healthcare professionals. In the future, we will develop more features to narrow the gap between the real world and the simulation environment such as simulating the deformable aorta, adding guidewire simulation, and considering the interaction of the instruments with the fluid. We will further bring our simulator to AR/VR environments.

REFERENCES

[1] I. Wamala, E. T. Roche, and F. A. Pigula, "The use of soft robotics in cardiovascular therapy," *Expert Review of Cardiovascular Therapy*, 2017.

[2] O. M. Omisore, S. Han, J. Xiong, H. Li, Z. Li, and L. Wang, "A review on flexible robotic systems for minimally invasive surgery," *SMC*, 2020.

[3] D. Kundrat, G. Dagnino, T. M. Kwok, M. E. Abdelaziz, W. Chi, A. Nguyen, C. Riga, and G.-Z. Yang, "An MR-Safe endovascular robotic platform: Design, control, and ex-vivo evaluation," *TBME*, 2021.

[4] V. M. Pereira, N. M. Cancelliere, P. Nicholson, I. Radovanovic, K. E. Drake, J.-M. Sungur, T. Krings, and A. Turk, "First-in-human, robotic-assisted neuroendovascular intervention," *Journal of Neurointerventional Surgery*, 2020.

[5] J. Burgner-Kahrs, D. C. Rucker, and H. Choset, "Continuum robots for medical applications: A survey," *T-RO*, 2015.

[6] T. Heidt, S. Reiss, A. J. Krafft, A. C. Özen, T. Lottner, C. Hehrlein, R. Galmbacher, G. Kayser, I. Hilgendorf, P. Stachon, *et al.*, "Real-time magnetic resonance imaging–guided coronary intervention in a porcine model," *Scientific Reports*, 2019.

[7] M. E. M. K. Abdelaziz, D. Kundrat, M. Pupillo, G. Dagnino, *et al.*, "Toward a Versatile Robotic Platform for Fluoroscopy and MRI-Guided Endovascular Interventions: A Pre-Clinical Study," in *IROS*, 2019.

[8] J. Konstantinova, A. Jiang, K. Althoefer, P. Dasgupta, and T. Nanayakkara, "Implementation of tactile sensing for palpation in robot-assisted minimally invasive surgery: A review," *IEEE Sensors Journal*, 2014.

[9] M. B. Molinero, G. Dagnino, J. Liu, W. Chi, M. E. Abdelaziz, T. M. Kwok, C. Riga, and G.-Z. Yang, "Haptic guidance for robot-assisted endovascular procedures: implementation and evaluation on surgical simulator," in *IROS*, 2019.

[10] A. Nguyen, D. Kundrat, G. Dagnino, W. Chi, M. E. Abdelaziz, Y. Guo, Y. Ma, T. M. Kwok, C. Riga, and G.-Z. Yang, "End-to-end real-time catheter segmentation with optical flow-guided warping during endovascular intervention," in *ICRA*, 2020.

[11] 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*, 2021.

[12] A. Juliani, V.-P. Berges, E. Teng, A. Cohen, J. Harper, C. Elion, C. Goy, Y. Gao, H. Henry, M. Mattar, *et al.*, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.

[13] L. Karstensen, T. Behr, T. P. Pusch, F. Mathis-Ullrich, and J. Stallkamp, "Autonomous guidewire navigation in a two dimensional vascular phantom," *CDBME*, 2020.

[14] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterlik, *et al.*, "SOFA: A Multi-Model Framework for Interactive Physical Simulation," in *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, 2012.

[15] R. Davis, R. Henshell, and G. Warburton, "A Timoshenko beam element," *JSV*, 1972.

[16] T. Behr, T. P. Pusch, M. Siegfarth, D. Hüsener, T. Mörschel, and L. Karstensen, "Deep Reinforcement Learning for the Navigation of Neurovascular Catheters," *Current Directions in Biomedical Engineering*, 2019.

[17] O. M. Omisore, T. Akinyemi, W. Duan, W. Du, and L. Wang, "A Novel Sample-efficient Deep Reinforcement Learning with Episodic Policy Transfer for PID-Based Control in Cardiac Catheterization Robots," *arXiv preprint arXiv:2110.14941*, 2021.

[18] E. Rohmer, S. P. Singh, and M. Freese, "V-REP: A versatile and scalable robot simulation framework," in *IROS*, 2013.

[19] P. Schegg, J. Dequidt, E. Coevoet, E. Leurent, R. Sabatier, P. Preux, and C. Duriez, "Automated Planning for Robotic Guidewire Navigation in the Coronary Arteries," in *RoboSoft*, 2022.

[20] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *IROS*, 2012.

[21] A. Attanasio, B. Scaglioni, E. De Momi, P. Fiorini, and P. Valdastri, "Autonomy in surgical robotics," *Control, Robotics, and Autonomous Systems*, 2021.

[22] W. Chi, G. Dagnino, T. M. Kwok, A. Nguyen, D. Kundrat, M. E. Abdelaziz, C. Riga, C. Bicknell, and G.-Z. Yang, "Collaborative robot-assisted endovascular catheterization with generative adversarial imitation learning," in *ICRA*, 2020.

[23] A. Marban, V. Srinivasan, W. Samek, J. Fernández, and A. Casals, "A recurrent convolutional neural network approach for sensorless force estimation in robotic surgery," *Biomedical Signal Processing and Control*, 2019.

[24] C. Gao, X. Liu, M. Peven, M. Unberath, and A. Reiter, "Learning to see forces: Surgical force prediction with rgb-point cloud temporal convolutional networks," in *MICCAI Workshop*, 2018.

[25] A. I. Aviles, S. M. Alsaleh, J. K. Hahn, and A. Casals, "Towards retrieving force feedback in robotic-assisted surgery: A supervised neuro-recurrent-vision approach," *ToH*, 2016.

[26] A. I. Aviles, S. M. Alsaleh, E. Montseny, P. Sobrevilla, and A. Casals, "A deep-neuro-fuzzy approach for estimating the interaction forces in robotic surgery," in *FUZZ-IEEE*, 2016.

[27] A. Marban, V. Srinivasan, W. Samek, J. Fernández, and A. Casals, "Estimation of interaction forces in robotic surgery using a semi-supervised deep neural network model," in *IROS*, 2018.

[28] C. I. Nesbitt, N. Birdi, S. Mafeld, and G. Stansby, "The Role of Simulation in the Development of Endovascular Surgical Skills," *Perspect Med Educ*, 2016.

[29] Y. Wei, S. Cotin, J. Dequidt, C. Duriez, J. Allard, E. Kerrien, *et al.*, "A (near) real-time simulation method of aneurysm coil embolization," *Aneurysm*, 2012.

[30] J. Dequidt, C. Duriez, S. Cotin, and E. Kerrien, "Towards interactive planning of coil embolization in brain aneurysms," in *MICCAI*, 2009.

[31] H. Talbot, F. Spadoni, C. Duriez, M. Sermesant, S. Cotin, and H. Delingette, "Interactive training system for interventional electro-cardiology procedures," in *ISBMS*, 2014.

[32] S. Sinceri, M. Carbone, M. Marconi, A. Moglia, M. Ferrari, and V. Ferrari, "Basic Endovascular Skills Trainer: A surgical simulator for the training of novice practitioners of endovascular procedures," in *EMBC*, 2015.

[33] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[34] J. Ho and S. Ermon, "Generative adversarial imitation learning," *NIPS*, 2016.

[35] Y. Cho, J.-H. Park, J. Choi, and D. E. Chang, "Image Processing Based Autonomous Guidewire Navigation in Percutaneous Coronary Intervention," in *ICCE-Asia*, 2021.

[36] J. Kweon, K. Kim, C. Lee, H. Kwon, J. Park, K. Song, Y. I. Kim, J. Park, I. Back, J.-H. Roh, *et al.*, "Deep reinforcement learning for guidewire navigation in coronary artery phantom," *IEEE Access*, 2021.

[37] G.-Z. Yang, J. Cambias, K. Cleary, E. Daimler, J. Drake, P. E. Dupont, N. Hata, P. Kazanzides, S. Martel, R. V. Patel, *et al.*, "Medical robotics–Regulatory, ethical, and legal considerations for increasing levels of autonomy," *Science Robotics*, 2017.

[38] M. Silcowitz, S. Niebe, and K. Erleben, "Interactive rigid body dynamics using a projected gauss–seidel subspace minimization method," in *Computer Vision, Imaging and Computer Graphics. Theory and Applications: International Joint Conference, VISIGRAPP 2010, Angers, France, May 17-21, 2010. Revised Selected Papers*, 2010.

[39] K. Mamou and F. Ghorbel, "A simple and efficient approach for 3D mesh approximate convex decomposition," in *ICIP*, 2009.

[40] G. Robinson and J. B. C. Davies, "Continuum robots-a state of the art," in *ICRA*, 1999.

[41] 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," *TBME*, 2016.

[42] K.-W. Kwok, K. H. Tsoi, V. Vitiello, J. Clark, G. C. Chow, W. Luk, and G.-Z. Yang, "Dimensionality reduction in controlling articulated snake robot for endoscopy under dynamic active constraints," *T-RO*, 2012.

[43] R. Featherstone, *Rigid body dynamics algorithms*, 2014.

[44] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[45] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *ICML*, 2018.

[46] G. Dalal, K. Dvijotham, M. Vecerik, T. Hester, C. Paduraru, and Y. Tassa, "Safe exploration in continuous action spaces," *arXiv preprint arXiv:1801.08757*, 2018.

[47] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.