

CathSim: An Open-source Simulator for Autonomous Cannulation

Tudor Jianu¹, Baoru Huang², Mohamed E. M. K. Abdelaziz², Minh Nhat Vu³, Sebastiano Fichera⁴, Chun-Yi Lee⁵, Pierre Berthet-Rayne⁶, Ferdinando Rodriguez y Baena⁷, Anh Nguyen¹

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 catheterisation 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 visualisation 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, radiation

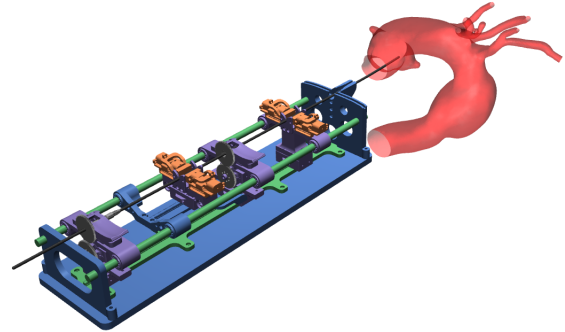


Fig. 1. An overview of our CathSim - an open-sourced simulator for autonomous cannulation. The figure shows the slave robot navigating the catheter through an aortic arch.

exposure to the surgeon, 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 master-slave teleoperation architecture have been proposed [3]–[5]. The surgeon actuates the master device from which, the information is mapped to the slave robot that executes the related action. The use of master-slave 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 ionising radiation exposure whilst allowing the soft tissue, such as the vasculature, to be visualised [6]. Furthermore, the robotic system provides additional information to the surgeon through assistive features such as force/torque information [7], [8], haptic feedback [9]–[14], and real-time segmentation and tracking [15], [16]. Nonetheless, the procedure continues to be manually conducted by the surgeon.

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 [26] and *ii*) increased duration of robotic procedure compared to its non-robotic counterpart [27]. Firstly, the surgeon operates within the tri-dimensional 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 extensive damage to the vasculature system thus operating under mentally strenuous conditions which can be diminished

¹Department of Computer Science, University of Liverpool, UK

²The Hamlyn Centre for Robotic Surgery, Imperial College London, UK

³Automation & Control Institute (ACIN), TU Wien, Austria

⁴Department of Mechanical, Materials & Aerospace Engineering, University of Liverpool, UK

⁵Department of Computer Science, National Tsing Hua University, Taiwan

⁶Honorary Fellow, University of Liverpool

⁷Mechatronics in Medicine Lab, Imperial College London, UK.

TABLE I
ENDOVASCULAR SIMULATORS COMPARISON

Simulator	Physics Engine	Catheter	Guidewire	Force Sensing	Open-source
Molinero <i>et al.</i> [12]	Unity Physics [17]	Discretised	✗	Vision-Based	✗
Karstensen <i>et al.</i> [18]	SOFA [19]	Timoshenko Beam theory [20]	✗	✗	✗
Behr <i>et al.</i> [21]	SOFA [19]	Timoshenko Beam theory [20]	✗	✗	✗
Omisore <i>et al.</i> [22]	CopelliaSim [23]	Unknown	✗	✗	✗
Schegg <i>et al.</i> [24]	SOFA [19]	Timoshenko Beam theory [20]	✗	✗	✗
CathSim (ours)	MuJoCo [25]	Discretised	✗	✓	✓

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

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 [26].

Recent developments in machine learning are promising a greater degree of autonomy in many robotic systems. Such systems leverage deep learning architectures such as convolutional neural networks [28], [29], long-short-term memory [30]–[32], and generative adversarial imitation learning [16] 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 [12], [18], [21], [22], [24], [27], [33]. 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 robotic autonomy through the provision of a standardised environment that confers familiarity to the machine learning community. As such, we propose a simulation environment that provides ease of interaction with a known architecture by using MuJoCo’s [25] physics engine. We employ two types of aortic arches with two different catheterisation tasks whilst using a discrete catheter with micro-motors-based actuation for increased control over the procedure. The overall model of our simulator is shown in Fig. 1. We summarise our contributions and the potential usage of CathSim 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 CathSim using two popular RL algorithms.
- 3) With real-time force sensing and high fidelity visualisation, 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

Endovascular Robotic Platforms. Recently, many robotic platforms have been proposed for endovascular intervention. Examples of such platforms are Magellan (AurisHealth, Redwood City, CA, USA), the R-one™ robot (Robocath, Rouen, France), the Amigo platform (Catheter Precision, Mt. Olive, NJ, USA), and the CorPath GRX platform (Corindus, A Siemens Healthineers Company, Waltham, MA, USA). Those platforms are designed for teleoperation and typically employ a master-slave architecture [34] which allows the surgeon to be situated outside of the operating room, thus diminishing the risk of repeated exposure to ionising radiation that the surgeon is facing. The surgeon manipulates the master robot which serves as an input to the slave robot. This information is then mapped to the end effector and visual-haptic feedback is provided back to the surgeon. Such platforms are designed for single-port-access and due to the dimensionality imposed by the confined spaces they operate in, the actuation is fairly limited. Furthermore, the space-compliant modelling of continuum catheters implies great challenges in terms of actuation control and contact handling.

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 [35] where each possesses its own limitations and advantages [36]–[38]. They mostly concentrate on the skill development of the trainee [35], [39] or the development and improvement of assistive features such as haptic feedback [12] or rely on physical materials which do not offer RL compliant environment. However, robotic learning research has been undertaken using synthetic simulators such as a high fidelity phantom [27] by means of imitation learning whereas other research used the SOFA [19] simulation environment to train a Deep Deterministic Policy Gradient (DDPG) [40] network with hindsight experience replay [41] 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 of collaborative improvement.

Autonomous Cannulation. The advancement of machine learning allows preliminary results of autonomous cannulation. In [27], generative adversarial imitation learning [42] has been used to perform different catheterisation tasks. The authors in [22] used deep RL in a simulated environment to

perform robotic axial navigation for coronary interventions. Cho *et al.* [43] created an image-based automatic control approach for controlling the guidewire. Kweon *et al.* [44] applied deep RL for guidewire navigation in a coronary artery phantom. For motion control during robotic catheterisation, Omisore *et al.* [22] used a sample-efficient deep RL with episodic policy transfer and a fully adaptive tuning technique. Within a two-dimensional phantom, Karstensen *et al.* [18] deployed a simulation-trained DDPG 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 [45] 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 suitable for training autonomous agents using different machine learning approaches. Based on MuJoCo’s [25] framework, CathSim offers an advanced simulation environment for real-time and faster than real-time applications. Furthermore, CathSim also provides real-time force sensing capability and high fidelity realistic visualisation 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. Based on its flexible design, we can also easily extend CathSim to other environments such as Augmented Reality (AR) and Virtual Reality (VR) applications.

III. PRELIMINARIES

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 (Fig. 2) follows the popular master-slave architecture that can be safely integrated into Magnetic Resonance (MR) settings due to its fabrication design [3] thus diminishing the exposure to ionising radiation whilst allowing soft tissue visualisation. Furthermore, the master robot uses haptic feedback generated by the catheter’s interaction with the environment through the navigation system [12], [46] and maintains an intuitive control that replicates human motion patterns such as insertion, retraction, and rotation. The slave robot mimics the master 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 complex mapping between the master and the slave robot in CathBot’s design [3], we focus only on simulating the slave robot for simplicity. The simulated model of the slave robot has been created through the decomposition of its CAD model into several sub-parts. The slave robot, depicted in Fig. 2, consists of 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

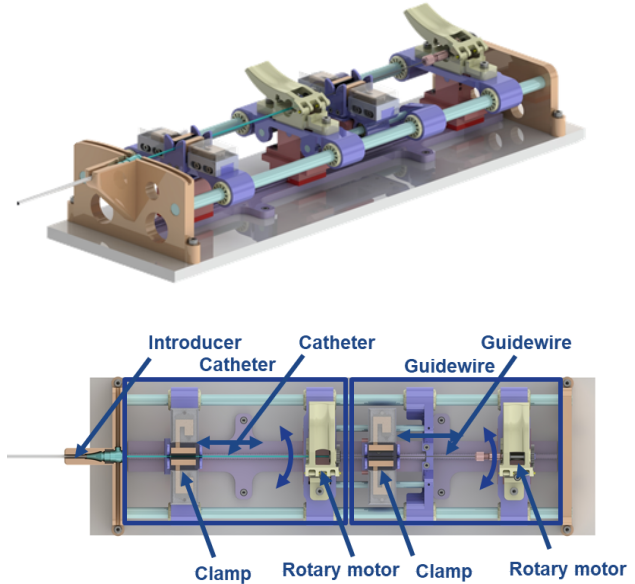


Fig. 2. The design of Cathbot’s slave robot. The two clamps lock the catheter in place by using pneumatic actuators. Those are situated on the main rail that facilitates the translational moves of the catheter. The figure also depicts the two rotary motors responsible for the rotational movements of the catheter and guidewire respectively.

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 clamps 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. As such, we actuate the catheter directly through the use of a prismatic joint. This method retains the control capabilities of the original CathBot design whilst facilitating manipulation learning rather than the underlying physics.

IV. SIMULATION AND LEARNING

A. Simulation Environment

Our CathSim environment is composed of three components: *i)* the slave robotic model for endovascular procedures [47], *ii)* the aortic arch phantoms, and *iii)* the catheter.

1) *Aorta:* 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 (V-HACD) [48] resulting in 1,024 convex hulls for Type-I aortic arch and 470 for Type-II aortic arch. The difference in the number of convex hulls is given by the difference in the measure of concavity of the two meshes which is explained in more detailed in [49]. Such collision models are

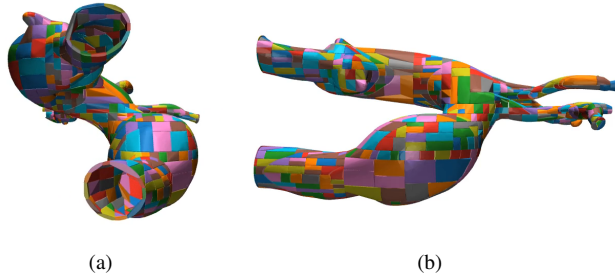


Fig. 3. The collision property of the aorta is inferred through the decomposition of the aortic arch into a series of convex hulls whilst the visual properties are given by scanned high fidelity silicone-based, transparent, anthropomorphic phantoms.

illustrated in Fig. 3, where the concatenation of the convex hulls generated from the Type-I aortic arch is presented. This representation aids the computational process and allows the use of soft contacts by the physics engine [25]. We include the original mesh without contact in order to maintain the visual realism of the system. The collision between the catheter and aorta is modelled as the point contacts between each convex hull of each element. In practice, the point contacts strategy can be solved rapidly in real-time using the convex optimisation problem through the Projected Gauss-Seidel method [48] while maintaining contact fidelity. Furthermore, albeit the rigidity of the convex hulls, the system allows penetrating contacts which imitate real-world dynamics.

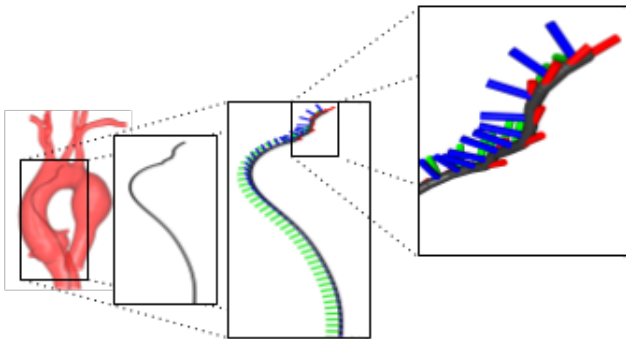


Fig. 4. The catheter’s continuous properties are approximated by a series of rigid bodies linked through revolute joints. The kinematic chain can be visualised through the frames of each individual body where the last ten links are actuated by internal micro-motors.

2) *Catheter*: One form of modelling the catheter represents the discretisation of the continuous shape into a series of rigid bodies interconnected by revolute or spherical joints [5]. This approach has been further categorised into discrete and serpentine methods, where the latter differs by having shorter rigid segments and, implicitly, a higher number of links [50]. The discretisation approach has been proven to confer reasonable accuracy for shape prediction of the continuum robot [51] where the authors paired the robot with Fibber Bragg Gratings (FBG) sensors. In contrast to the discretisation approach, the continuous curve estimation methods [5] assume that the described object is composed of an elastic backbone where the model aims to describe the

shape of the underlying 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, we model the catheter based on a discretisation approach, creating a serpentine-like model in which its continuous shape and deformation are approximated by a series of 100 rigid bodies joined together by two revolute joints, resulting in 198 revolute joints. This configuration transfers into a catheter length of 72 cm whereas an endovascular-based catheter has a length of 65-125 cm. A prismatic joint was added for mimicking the direct linear movement of the catheter resulting in an 199 degrees of freedom system. We limit the actuation to the tip of the catheter, represented by the last 10 links or 7 cm, where the revolute joints are intrinsically actuated by motors, similarly to [52]. This approach allows the use of standardised modelling which is a good approximation of the continuous robot whilst benefiting from an increased actuation accuracy compared to its extrinsic actuated continuous counterparts. The catheter model is shown in Fig. 4.

B. Reinforcement Learning

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 timestep 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 $g \in \mathcal{G}$.

1) *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. Given the degrees of freedom of the system $n_V = 199$ and the number of bodies $n_{bodies} = 100$, we provide our system with the following information:

- position: $n_V \times 1$
- velocity: $n_V \times 1$
- centre of mass inertia: $n_{bodies} \times 10$
- centre of mass velocity: $n_{bodies} \times 6$
- actuators generated force: $n_V \times 1$
- external forces generated on body: $n_{bodies} \times 6$

Secondly, we use the Image observation by positioning an RGB camera of $128 \times 128 \times 3$ resolution on top of the aortic phantom. We map the image from RGB to grayscale as in [39], 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.

2) *Actions*: 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 normalised within a $[-1, 1]$ interval, giving a space of $a_t \in [-1, 1]^{21}$.

3) *Rewards*: 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.

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

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.

4) *Network Architectures*: Considering the continuous action representation, we employ two state-of-the-art RL algorithms, namely PPO [53] and SAC [54] with the parameters utilised in [55]. 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 [56] and use CNN for both Image and Sequential observations as they only have one and four input channels respectively where each channel represents a grayscale 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 optimiser is used to train all networks with a learning rate of 0.0003.

V. EXPERIMENTS

A. Experiment Setup

We consider the autonomous catheterisation 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

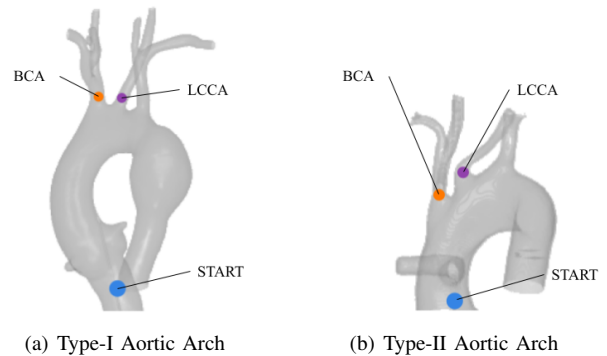


Fig. 5. The starting configuration. The figure depicts the navigation task employed in the Type-I and Type-II Aortic Arches. The catheter is initially situated within the ascending aorta with the task of navigating towards the brachiocephalic artery (BCA) or the left common carotid artery (LCCA). The task finishes when the tip of the catheter is situated within a proximity of 8 mm of the targets.

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. 5).

For each of the aforementioned tasks, we train the model based on the setup from Section IV-B. for a total of 600,000 timesteps. Within each episode, the catheter is randomly initialised with a displacement of 1 mm. The agent proceeds the navigation throughout the ascending aorta until it reaches the goal state or upon reaching 2,000 steps. At each step, we extract the contact points between the aorta along with the force and project it on a 2D image that overlaps the image captured by the RGB camera situated on top of the aorta where each pixel is associated with the value of the force resulting from the contact. The overlapping of the two images results in the heatmap in Fig. 6. We proceed with an evaluation undertaken on 30 episodes, extracting the force interaction and the reward information for each episode. The average results are summarised in Table II whilst the force distribution can be visualised within Fig. 7. The training time varies 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.

B. Quantitative Results

Using the extracted force images for each timestep, 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 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 although it shows the least performance when the target is LCCA. The success rate mimics the reward, where the highest success rate resulted from the former and no success in the latter. A more coherent

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	-140 ± 96	0.005 ± 0.007	0.007 ± 0.016	0.048 ± 0.020	0.076 ± 0.059	83	37
	SAC-Image	-391 ± 157	-196 ± 5	0.024 ± 0.035	0.003 ± 0.003	0.258 ± 0.094	0.059 ± 0.028	00	00
	PPO-Sequential	-57 ± 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	0 ± 0	-67 ± 30	0.014 ± 0.010	0.010 ± 0.016	0.030 ± 0.007	0.121 ± 0.044	100	27
	PPO-Image	-20 ± 47	-8 ± 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	0.004 ± 0.005	0.005 ± 0.005	0.042 ± 0.013	0.050 ± 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

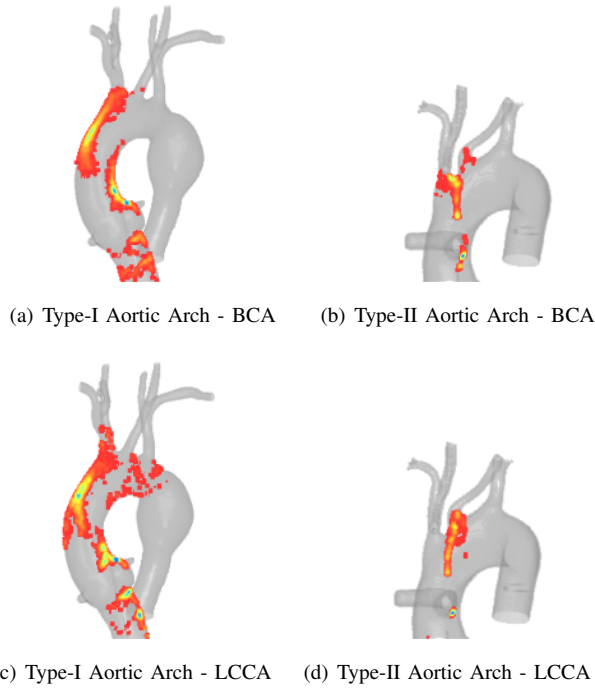


Fig. 6. The force heatmap extracted from the interaction between the catheter and the aortas when undertaking the navigating towards the BCA target on the top row, and LCCA target on the second row. The catheterisation procedure exerts a greater force on the first aorta where complicated manoeuvres are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

reward has been achieved whilst using a singular image observation, where the cannulation of LCCA presents the greatest success. In contrast, the cannulation of BCA presents close results to the sequential observation. The higher reward translates into shorter cannulation and therefore a lower risk for the patient. The catheterisation 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 catheterisation procedure. Overall, PPO shows better suitability for the task, especially when the observation is image-based.

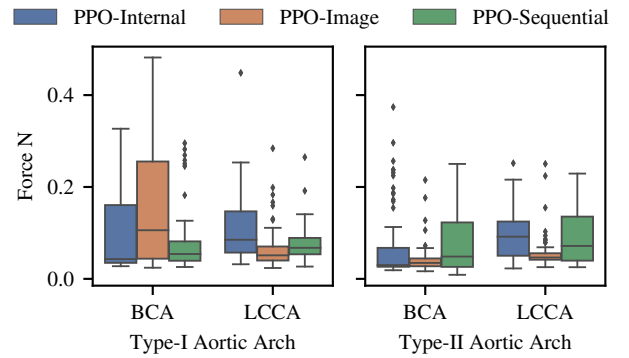


Fig. 7. The max forces generated by the catheter on the aorta within each of the evaluation episodes during the PPO training are shown. The left figure depicts the results gathered within the Type-I Aortic Arch where the catheter has greater interaction with the aorta, whereas the one on the right shows the results gathered within the Type-II Aortic Arch where the navigation is more direct.

C. Qualitative Result

We use the force frames extracted from each timestep 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.

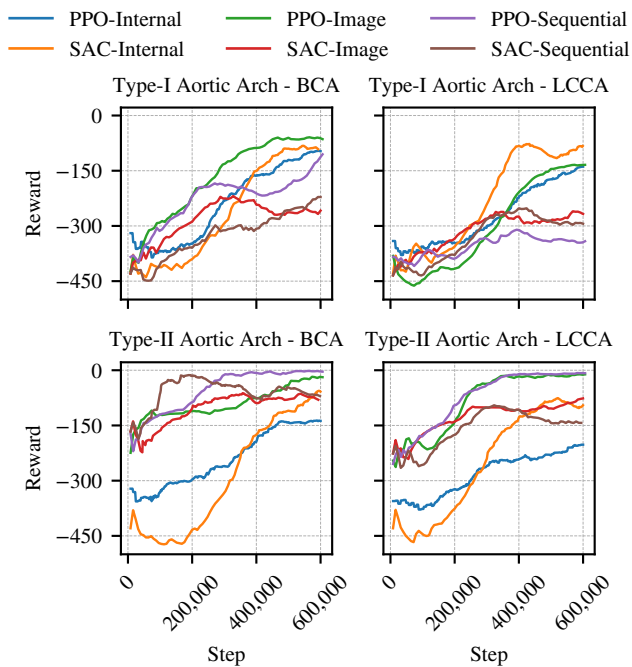


Fig. 8. The rewards when training autonomous agents using an on-policy algorithm (PPO) and an off-policy algorithm (SAC). In general, we observe that PPO performs better than SAC. The image observation consistently achieved good results when using PPO. Furthermore, the sequential observation shows the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

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 visualised 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 the environments, PPO with an image observation obtained the highest or close to highest reward, thus showing that an on-policy algorithm, performs generally 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 or in other words, it is capable of reaching the target in the lowest amount of time. 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 benchmark-

ing 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 make our simulator close to a real-world setting, including the deformable aorta, simulating the guidewire, and further bringing 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*, vol. 15, no. 10, pp. 767–774, 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," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 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," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 10, pp. 3110–3121, 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*, vol. 12, no. 4, pp. 338–340, 2020.
- [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] 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*, vol. 9, no. 1, pp. 1–10, 2019.
- [7] 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*, vol. 14, no. 8, pp. 2490–2501, 2014.
- [8] D. G. Black, A. H. H. Hosseinabadi, and S. E. Salcudean, "6-dof force sensing for the master tool manipulator of the da vinci surgical system," *IEEE Robotics and automation letters*, 2020.
- [9] H. Naghibi, W. B. Hoitzing, S. Stramigioli, and M. Abayazid, "A flexible endoscopic sensing module for force haptic feedback integration," in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*. IEEE, 2018, pp. 158–161.
- [10] K. Ogawa, M. Y. Ibrahim, and K. Ohnishi, "Development of flexible haptic forceps based on the electrohydraulic transmission system," *IEEE Transactions on Industrial Informatics*, 2018.
- [11] R. Moccia, M. Selvaggio, L. Villani, B. Siciliano, and F. Ficuciello, "Vision-based virtual fixtures generation for robotic-assisted polyp dissection procedures," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.
- [12] 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*. IEEE, 2019.
- [13] O. A. Van der Meijden and M. P. Schijven, "The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review," *Surgical endoscopy*, vol. 23, no. 6, pp. 1180–1190, 2009.
- [14] A. Spinelli, G. David, S. Gidaro, M. Carvello, M. Sacchi, M. Montorsi, and I. Montroni, "First experience in colorectal surgery with a new robotic platform with haptic feedback," *Colorectal Disease*, 2018.
- [15] F. Alambeigi, Z. Wang, R. Hegeman, Y.-H. Liu, and M. Armand, "Autonomous data-driven manipulation of unknown anisotropic deformable tissues using unmodelled continuum manipulators," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 254–261, 2018.
- [16] 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 *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 9967–9973.

- [17] 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.
- [18] L. Karstensen, T. Behr, T. P. Pusch, F. Mathis-Ullrich, and J. Stalkamp, “Autonomous guidewire navigation in a two dimensional vascular phantom,” *Current Directions in Biomedical Engineering*, vol. 6, no. 1, 2020.
- [19] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterlik, and S. Cotin, “Sofa: A multi-model framework for interactive physical simulation,” in *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, 2012. [Online]. Available: <https://hal.inria.fr/hal-00681539>
- [20] R. Davis, R. Henshell, and G. Warburton, “A timoshenko beam element,” *Journal of Sound and Vibration*, 1972.
- [21] 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*, vol. 5, no. 1, pp. 5–8, Sept. 2019.
- [22] 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.
- [23] E. Rohmer, S. P. Singh, and M. Freese, “V-rep: A versatile and scalable robot simulation framework,” in *IROS*, 2013.
- [24] 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 *2022 IEEE 5th International Conference on Soft Robotics (RoboSoft)*. IEEE, 2022, pp. 239–246.
- [25] E. Todorov, T. Erez, and Y. Tassa, “Mujoco: A physics engine for model-based control,” in *IROS*, 2012.
- [26] A. Attanasio, B. Scaglioni, E. De Momi, P. Fiorini, and P. Valdastrì, “Autonomy in surgical robotics,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 651–679, 2021.
- [27] 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 *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 2414–2420.
- [28] 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*, vol. 50, pp. 134–150, 2019.
- [29] 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 *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. Springer, 2018, pp. 118–127.
- [30] 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,” *IEEE transactions on haptics*, vol. 10, no. 3, pp. 431–443, 2016.
- [31] 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 *2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2016, pp. 1113–1119.
- [32] 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 *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [33] G. Fagogenis, M. Mencattelli, Z. Machaidze, B. Rosa, K. Price, F. Wu, V. Weixler, M. Saeed, J. E. Mayer, and P. E. Dupont, “Autonomous robotic intracardiac catheter navigation using haptic vision,” *Science robotics*, vol. 4, no. 29, 2019.
- [34] R. D. Howe and Y. Matsuoka, “Robotics for surgery,” *Annual review of biomedical engineering*, vol. 1, no. 1, pp. 211–240, 1999.
- [35] C. I. Nesbitt, N. Birdi, S. Mafeld, and G. Stansby, “The role of simulation in the development of endovascular surgical skills,” *Perspect Med Educ*, vol. 5, no. 1, pp. 8–14, Feb. 2016.
- [36] 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*, vol. 8, no. 29, pp. 223–248, 2012.
- [37] H. Talbot, F. Spadoni, C. Duriez, M. Sermesant, S. Cotin, and H. Delingette, “Interactive training system for interventional electrocardiology procedures,” in *International Symposium on Biomedical Simulation*. Springer, 2014, pp. 11–19.
- [38] 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 *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015.
- [39] 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.
- [40] M. Andrychowicz, F. Wolski, A. Ray, J. Schneider, R. Fong, P. Welinder, B. McGrew, J. Tobin, O. Pieter Abbeel, and W. Zaremba, “Hindsight experience replay,” *Advances in neural information processing systems*, vol. 30, 2017.
- [41] J. Ho and S. Ermon, “Generative adversarial imitation learning,” *Advances in neural information processing systems*, vol. 29, 2016.
- [42] Y. Cho, J.-H. Park, J. Choi, and D. E. Chang, “Image processing based autonomous guidewire navigation in percutaneous coronary intervention,” in *2021 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*. IEEE, 2021, pp. 1–6.
- [43] 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.
- [44] 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*, vol. 2, no. 4, p. 8638, 2017.
- [45] G. Dagnino, J. Liu, M. E. Abdelaziz, W. Chi, C. Riga, and G.-Z. Yang, “Haptic feedback and dynamic active constraints for robot-assisted endovascular catheterization,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018.
- [46] 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 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.
- [47] M. Silcowitz, S. Niebe, and K. Erleben, “Interactive rigid body dynamics using a projected gauss–seidel subspace minimization method,” in *International Conference on Computer Vision, Imaging and Computer Graphics*. Springer, 2010, pp. 218–229.
- [48] K. Mamou and F. Ghorbel, “A simple and efficient approach for 3d mesh approximate convex decomposition,” in *2009 16th IEEE international conference on image processing (ICIP)*, 2009.
- [49] G. Robinson and J. B. C. Davies, “Continuum robots—a state of the art,” in *Proceedings 1999 IEEE international conference on robotics and automation*, 1999.
- [50] 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, 2016.
- [51] 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,” *IEEE Transactions on robotics*, vol. 29, no. 1, pp. 15–31, 2012.
- [52] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [53] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in *International conference on machine learning*, 2018.
- [54] 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.
- [55] 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.