From Patient-specific Digital Twin to Real-world Phantom: Autonomous Right Heart Catheterization

Yaxi Wang[1], *Student member, IEEE*, Mengzhe Xu[2], Wenlong Gaozhang[1] and Helge A Wurdemann[1], *Member, IEEE*

***Abstract*—Right heart catheterization (RHC) is a critical procedure for diagnosing and managing cardiovascular diseases (CVDs) such as heart failure, congenital heart disease, pulmonary edema, and pulmonary hypertension. However, currently prevalent manual RHC procedures requires continuous communication of clinicians between the main control room and the operating room, leading to navigation inaccuracies and increased physical workload for clinicians during prolonged procedure. To overcome these challenges, this paper introduces a robotic system that enables autonomous RHC (Auto-RHC) by transferring a catheter decision-making model from patient-specific digital twins to real-world robotic intervention using deep learning (DL) algorithms. By creating a patient-specific (PS) digital twin using the Simulation Open Framework Architecture (SOFA) and conducting virtual RHC interventions, images capturing the catheter balloon's position and aligned behavioral datasets were collected and utilized as input for a convolutional neural network (CNN) architecture. The trained catheter decision-making model derived from the digital twin was then transferred to real-world implementations of robot-assisted Auto-RHC. Experimental results validated the performance of the digital twin and demonstrated that the real-world robotic Auto-RHC achieved a high success rate across both static ($\geq$ 96%) and dynamic heartbeat ($\geq$ 94%) patient-specific cardiac phantoms. Furthermore, Auto-RHC enhanced navigation consistency by $\geq$ 34.63% compared to expert manual operation.**

***Index Terms*—Cardiovascular disease, sim-to-real, interventional robotics, machine learning, right heart catheterization.**

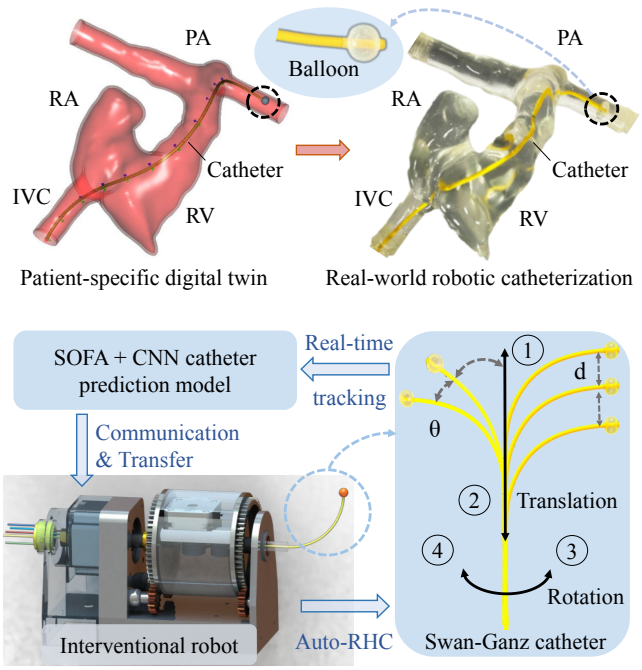


Fig. 1. Transferring from digital twin to real-world robotic RHC. Data incorporating the behaviour and position of the balloon was collected via the digital twin as input to train the prediction model. Then communication was established with the robot, using the 3D real-time balloon tracking as feedback to autonomously operate the catheter for RHC procedure.

I. Introduction

CARDIOVASCULAR diseases (CVDs) are the leading cause of mortality worldwide, with the number of deaths increasing each year [1]. In 2021, CVDs accounted for 19.8 million deaths globally, representing approximately one-third of all mortalities. Common symptoms CVDs include congenital heart disease, shock, heart failure, and pulmonary hypertension. Right heart catheterization (RHC), also known as pulmonary artery catheterization (PAC), is a key diagnostic and therapeutic procedure for managing these conditions [2]. Current clinical RHC procedures rely on real-time computed tomography (CT) or magnetic resonance imaging (MRI), with clinicians manually operating a Swan-Ganz catheter. The pre-curved, flexible catheter is inserted through the femoral, jugular or subclavian vein and navigated through the endovascular, advancing from caval vein (CV) to the right atrium (RA), right ventricle (RV), and ultimately into the pulmonary artery (PA) for diagnosis and treatment. However, manual RHC procedure introduces latency and operator-dependent variability, which diminishes the precision and consistency of catheter balloon navigation within the patient's heart.

To address the challenges of manual endovascular and cardiovascular interventions, a number of intelligent robotic systems incorporating machine learning (ML) techniques have been developed, demonstrating the potential for implementing autonomous interventions. ML-based interventional robotics typically involve learning from demonstration (LfD) and simulation-to-real-world (sim-to-real) transfer using deep learning (DL) or reinforcement learning (RL) algorithms [11]–[13]. Notably, H. Rafii-Tari et al. developed an LfD framework using a Gaussian mixture model (GMM) to replicate expert catheter manipulation by generating smooth trajectories, which was validated within phantom environment [14]. Building on

This work is supported by the Springboard Award of the Academy of Medical Sciences (grant number: SBF003-1109), the Engineering and Physical Sciences Research Council (grant numbers: EP/R037795/1, EP/S014039/1 and EP/V01062X/1), The UCL Dean's Prize, UCL Mechanical Engineering

[1]Yaxi Wang, Wenlong Gaozhang and Helge A. Wurdemann are with the Department of Mechanical Engineering, University College London, UK. h.wurdemann@ucl.ac.uk

[2]Mengzhe Xu is with the Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

TABLE I
SUMMARY OF RELATED SIM-TO-REAL AUTONOMOUS CARDIOVASCULAR/ENDOVASCULAR INTERVENTIONS

Ref. and Key authors	Sim-to-real methodology	Navigation type	Validation environment	PS simulation	PS validation
[3] O. Omisore. et al. (2021)	CopelliaSim using DRL-PID	3D catheter	In virtual static phantom	N/A	N/A
[4] Y. Cho et al. (2022)	SOFA using DDPG	2D guidewire	In vitro static phantom	✗	✗
[5] L. karstensen et al. (2022)	SOFA using DDPG	2D guidewire	In vitro static phantom	✔	✔
[6] Y. Wang et al. (2022)	VR simulator using DDPG	3D catheter	In virtual static phantom	✔	N/A
[7] Z. Mei et al. (2024)	Pymunk using PPO	2D guidewire	In vitro static phantom	✗	✗
[8] V. Scarponi et al. (2024)	SOFA using Zero-shot	3D guidewire	In vitrual static phantom	✗	N/A
[9] H. Robertshaw et al. (2024)	SOFA using IRL	3D catheter/guidewire	In vitrual static phantom	✔	N/A
[10] T. Jianu et al. (2024)	CathSim using ENN	3D catheter	In vitro static phantom	✗	✗
This Work	SOFA using CNN	3D catheter	In vitro dynamic/static patient-specific	✔	✔

this, W. Chi et al. enhanced the LfD approach by modeling proximal catheter movement with GMM, achieving a 98% success rate and a 33.3% reduction in mean contact force in an aortic arch phantom [15]. However, LfD typically require extensive demonstrations data from physical environments, which limits their practicality in clinical applications. To overcome these limitations, sim-to-real transfer learning has emerged as a promising solution [16].

Table I summarizes the state-of-the-art sim-to-real learning-based cardiovascular/endovascular interventional robotics, including learning methodologies, navigation types, validation environments, and simulation/validation fidelity. In this study, patient-specific (PS) simulation denotes to virtual interventions based on patient-specific anatomical models. The PS phantom refers to in vitro patient-specific models for experimental validation. For example, O. Omisore et al. developed a robot-assisted percutaneous coronary interventions (PCI) simulation platform using CoppeliaSim (Coppelia Robotics), integrating a deep reinforcement learning (DRL) framework with a PID control strategy [3]. While the study demonstrated sample-efficient transfer learning, validation using virtual or in-vitro phantoms was not conducted. To address the sim-to-real transfer gap, Y. Cho et al. established a SOFA platform for PCI [4]. By integrating behavioral cloning (BC) with the deep deterministic policy gradient (DDPG) algorithm, their neural network-enabled model achieved autonomous 2D guidewire navigation in an in-vitro vascular phantom. Extending this approach, L. Karstensen et al. applied a similar SOFA-DDPG methodology to an in vitro porcine liver-specific model. While the model successfully enabled 2D autonomous navigation, the success rate declined by 30% compared to the 100% success rate achieved in simulation. Y. Wang et al. developed an RL-based virtual reality (VR) training system, transferring expert navigation skills to novice operators and minimizing vascular collision risks during endovascular procedures [6]. Z. Mei et al. employed the Proximal Policy Optimization (PPO) algorithm, demonstrating successful sim-to-real transfer for guidewire navigation, with enhanced performance in real-world interventions [7]. Further contributions include V. Scarponi et al., who utilized a zero-shot learning methodology within the SOFA for 3D guidewire navigation, eliminating the need for retraining during real-world deployment. Similarly, H. Robertshaw et al. applied inverse reinforcement learning (IRL) for 3D catheter/ guidewire navigation, enhancing adaptability in static environments. Furthermore, T. Jianu et al. introduced the open-source CathSim simulator, combining an interventional robotic system with trained model to replicate expert trajectories for vascular interventions [10]. While the projects summarized in Table I have made significant advancements in sim-to-real autonomous endovascular and cardiovascular navigation, achieving patient-specific simulation and implementing autonomous interventions in dynamic, patient-specific anatomical environments remain critical challenges for real-world robotic clinical applications.

The contribution of this paper is to transfer the SOFA-built patient-specific digital twin integrated with CNN architecture, to real-world robotic autonomous right heart interventions in vitro static/dynamic patient-specific phantoms. The RHC digital twin was created within the SOFA framework, allowing user-controlled navigation of a Swan-Ganz balloon catheter to interact with phantoms from IVC to PA. Behavioral and positional data of the calibrated balloon were collected during repeated virtual procedures. This dataset served as input for training a catheter behaviour prediction model using the CNN architecture. In the real-world environment, communication was established between the robot controller, trained model and balloon tracking system to enable autonomous catheter navigation. Experimental validation demonstrated high Auto-RHC success rates across various patient-specific scenarios, achieving $\geq 96\%$ in static environments and $\geq 94\%$ in dynamic pulsatile heartbeat conditions. Notably, Auto-RHC shows a significant improvement ($\geq 34.63\%$) in trajectory consistency compared to clinical expert manual operations.

Section II details the Swan-Ganz catheter modelling, creation of the RHC digital twin, together with the CNN architecture and training of the catheter decision-making model. The validation experiments from digital twins to the real-world Auto-RHC in three scenarios are presented in Section III. Section IV concludes this study and future work.

II. DIGITAL TWIN AND ROBOT PREDICTION MODEL

A. Modelling of the Swan-Ganz catheter

Creating a digital twin of the Swan-Ganz catheter is challenging due to its flexible body ($L = 1000\,\text{mm}$), pre-curved tip ($R = 70\,\text{mm}$), and inflatable balloon ($D = 8\,\text{mm}$). This study incorporated the physical Swan-Ganz catheter model and

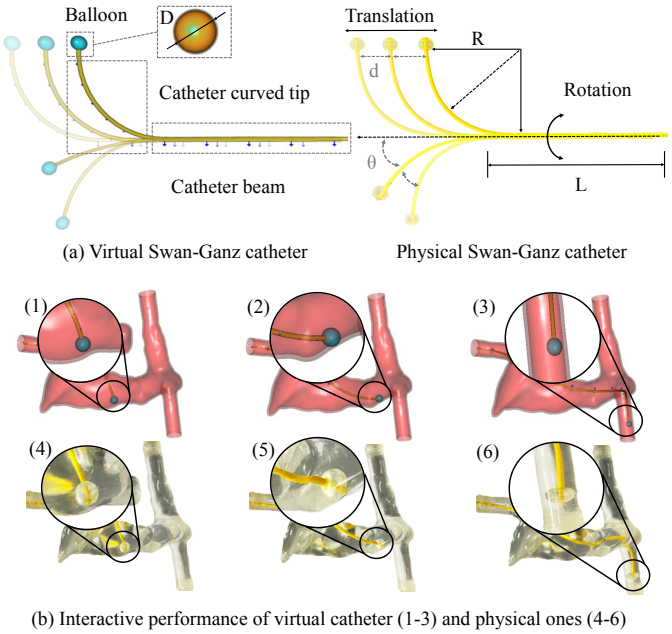


Fig. 2. Swan-Ganz catheter modelling and its performance of interactions. (a) shows the virtual catheter and physical Swan-Ganz catheter. It consists of a balloon, pre-curved tip and body (b) Comparison of the performance of catheter interacting with phantom in digital twin (1-3) and real-world (4-6).

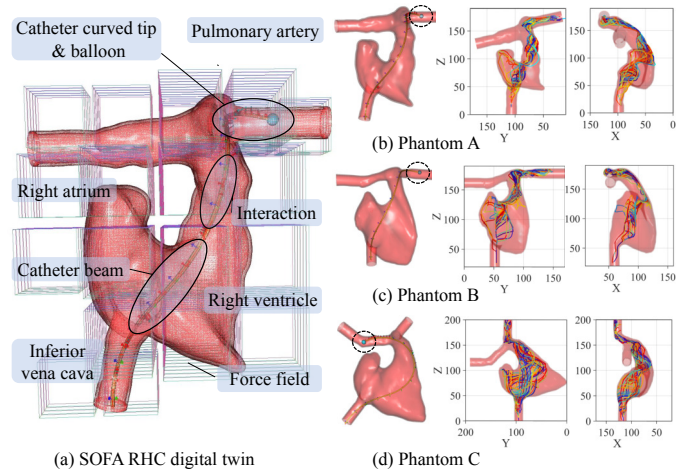


Fig. 3. RHC digital twin and data collection based on SOFA framework. (a) RHC digital twin includes a Swan-Ganz catheter and patient-specific phantom. (b) Virtual RHC intervention for patient-specific phantom A and 50 repetitions to record catheter balloon position and behaviours dataset including the X-Z and Y-Z plane. (c) Virtual intervention of phantom B and data collection. (d) Virtual intervention of phantom C and data collection.

Kirchhoff rod theory in the SOFA v22.12 framework (SOFA-BeamAdapter) to build the catheter digital twin [17]. The digital twin scene was created in SOFA-Python3. Catheter modelling in the SOFA framework provides models of visualization, collision and interaction (See the supplementary video). The deformation of the catheter is characterized by the displacement $\vec{u}(s)$ and twist angle $\phi(s)$, both derived from the total energy E_{total}, which comprises the sum of energies resulting from bending, twisting, and external forces:

$$E_{\text{total}} = E_{\text{bend}} + E_{\text{twist}} + E_{\text{external}} \tag{1}$$

The bending energy E_{bend} can be expressed as:

$$E_{\text{bend}} = \frac{1}{2}\int_0^L EI(s)\kappa_b(s)^2\,ds \tag{2}$$

Where $EI(s)$ is bending stiffness of the catheter, $\kappa(s)$ is curvature at position s along its length, and L is the length of the catheter. The twisting energy E_{twist} is:

$$E_{\text{twist}} = \frac{1}{2}\int_0^L GJ(s)\left(\frac{d\phi(s)}{ds}\right)^2 ds \tag{3}$$

Where $GJ(s)$ is the torsional stiffness, $\phi(s)$ is the twist angle as a function of s.

$$\begin{aligned} E_{\text{external}} = &-\left(\vec{F}_{\text{push}}\cdot\vec{u}(s_{\text{entrance}}(t)) + T_{\text{twist}}\cdot\phi(s_{\text{entrance}}(t))\right) \\ &-\int_0^L \vec{F}_{\text{contact}}(s)\cdot\vec{u}(s)\,ds \end{aligned} \tag{4}$$

This equation elaborates how external forces—push ($\vec{F}_{\text{push}}$) and twist (T_{twist})—are dynamically applied at the entrance point $s_{\text{entrance}}(t)$, which varies over time to reflect the dynamic nature of catheter insertion. $\vec{F}_{\text{contact}}(s)$ is the distributed contact forces along the catheter.

$$\begin{cases} EI(s)\dfrac{d^4\vec{u}(s)}{ds^4} + \vec{F}_{\text{push}}\cdot\delta(s - s_{\text{entrance}}(t)) - \vec{F}_{\text{contact}}(s) = \vec{0} \\ GJ(s)\dfrac{d^2\phi(s)}{ds^2} + T_{\text{twist}}\cdot\delta(s - s_{\text{entrance}}(t)) - \tau_{\text{contact}}(s) = 0 \end{cases} \tag{5}$$

Here, $\delta(s - s_{\text{entrance}}(t))$ localizes the application of external forces and torques at the entrance ($s = s_{\text{entrance}}(t)$). $\tau_{\text{contact}}(s)$ represents the torque due to contact forces along the catheter.

Contact detection between the catheter and the patient-specific phantom is implemented using SOFA's built-in collision detection module. A penalty-based method is employed to compute the distributed contact force $\vec{F}_{\text{contact}}(s)$ and $\tau_{\text{contact}}(s)$. The governing equations derived from Kirchhoff rod theory are discretized using SOFA's BeamAdapter module, which provides a computationally reliable framework for simulating flexible structures. The system is solved numerically using the EulerImplicitSolver, which employs a backward Euler time-stepping scheme to ensure accuracy in simulating the catheter's deformation. The digital phantom was assigned uniform properties to enable real-time interaction: Young's modulus of 10 kPa, Poisson's ratio of 0.49, density of 1.2 g/cm^3 and Rayleigh damping coefficients $\alpha = 0.1$, $\beta = 0.1$. The catheter mesh model is built programmatically in SOFA-Python3, featuring 200 edges and a spire diameter of ($R = 70$ mm), which accurately represents the pre-curved catheter.

B. Digital twins for RHC procedure

Fig. 3(a) shows the RHC digital twin created under the SOFA framework employing patient specific phantom. It begins with the initialization of the Catheter-Balloon Controller, followed by acquiring the catheter and balloon nodes. Event

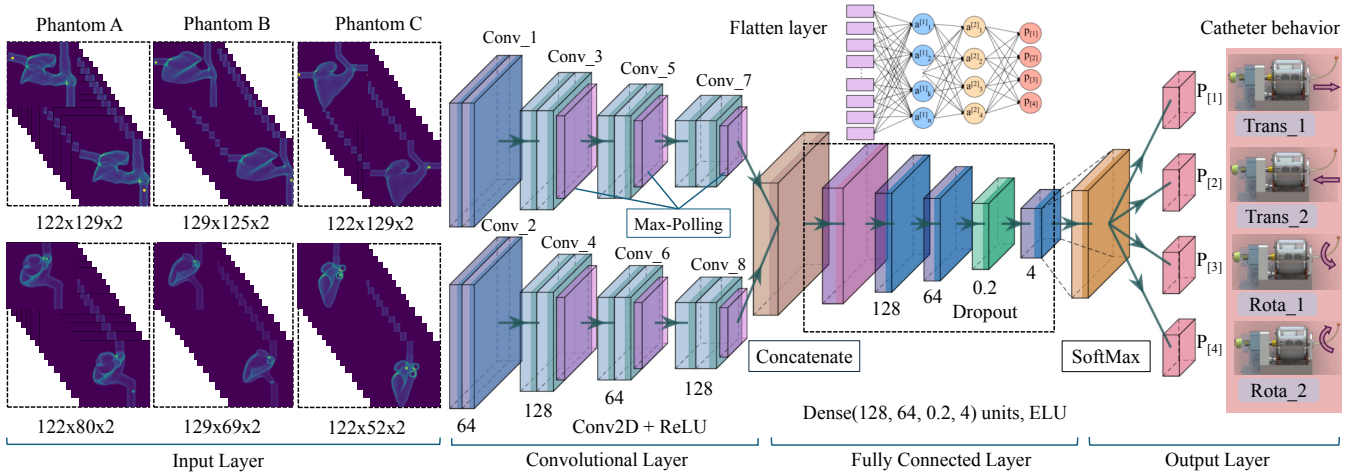


Fig. 4. CNN architecture and policy training. The DL-based CNN for training robot decision-making policies consists of input layer, convolutional layer, fully connected layer and output layer. The input layer is the data from SOFA RHC digital twin and output layer presents behavioural prediction of the robot.

handling functions are then implemented to manage animation and keypress event, with a Key-Pressed Controller developed in SOFA-Python3 to enable the catheter's translation and rotation. Next, the simulation environment is built by integrating essential plugins and models, such as Beam-Adapter and Edge-Topology. A patient-specific cardiac phantom is introduced, equipped with solvers of the Euler-Implicit Solver and Hexahedron FEM Force Field to simulate its mechanical behavior. The simulation is executed with the SOFA runtime and GUI modules initializing the scene. The catheter balloon is tracked and its 3D position and corresponding actions are recorded during the intervention. As is shown in Fig. 3(b), each phantom undergoes intervention procedure 50 times.

Fig. 2(b) presents the validation of the replicability of the RHC digital twin by comparing it to real-world execution. In the RHC digital twin, the position of the catheter balloon was tracked while controlling the catheter's interaction with the phantom at a sampling frequency of 40 HZ. The catheter's movement was defined by discrete actions: 1: translation forward, 2: translation backward, 3: rotation right, and 4: rotation left, where each step involved a 20° rotation and a 5mm translation. When the balloon reached the predefined PA, the output action is "0", indicating task completion. For the physical RHC, the catheter balloon was calibrated using an NDI Aurora 6-DoF sensor (Aurora, NDI, Canada) (dimensions: $L \times D = 9$ mm $\times 1.5$ mm), and catheter manipulation was performed via a 2-DOF robotic system (Fig. 1, bottom left) controlled using a joystick. The tracking system frequency was also set to 40 HZ, ensuring a time-series-consistent 3D trajectory dataset in both virtual and real-world environments.

To quantitatively assess the replicability of the RHC digital twin in intra-phantom procedures, this study employed Mean Euclidean Distance (MED), Maximum Deviation $D_{\max}$ and Root Mean Square Error E_{rms} to compare the consistency between virtual and real catheter trajectories. Given two synchronized 3D trajectories from the virtual and real environments, where both have the same sampling frequency, movement parameters, and total duration.

The virtual trajectory T_{sim} consists of a sequence of 3D coordinate points $(x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_n, y_n, z_n)$ while the real-world trajectory T_{real} is described by the corresponding sequence of 3D coordinates $(x'_1, y'_1, z'_1), (x'_2, y'_2, z'_2), \ldots, (x'_n, y'_n, z'_n)$ where n denotes the total number of sampled points, the data points recorded in this experiment are 600.

Mean Euclidean Distance (MED) is the average deviation between corresponding points is computed as:

$$MED = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2 + (z_i - z'_i)^2} \quad (6)$$

Maximum Deviation ($D_{\max}$ is the maximum discrepancy between corresponding points is given by:

$$D_{\max} = \max_{1 \le i \le n}\sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2 + (z_i - z'_i)^2} \quad (7)$$

E_{rms} quantifies the overall deviation as:

$$E_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[(x_i - x'_i)^2 + (y_i - y'_i)^2 + (z_i - z'_i)^2\right]} \quad (8)$$

Using the collected 3D datasets of virtual and real-world robot-operated catheter balloon trajectories within the phantom as inputs, the values of MED, $D_{\max}$ and E_{rms} can be calculated. The results show that the MED is 2.31 mm, the $D_{\max}$ is 4.63 mm, the E_{rms} is 2.74 mm. These values are well within clinically acceptable tolerances for catheter navigation procedures, as positional deviations under 5 mm are generally considered non-critical and do not impact procedural outcomes [18]. It is demonstrated that the SOFA-based RHC digital twin is highly replicable for real-world catheter interventions on patient-specific phantoms.

C. CNN architecture and policy training

Fig. 4 illustrates the CNN architecture for catheter decision-making prediction model training. It consists of input layer, convolutional layers, fully connected layers and output layer. The input layer receives stereo image sequences from the SOFA digital twin, capturing the catheter's top and side views within the vascular phantom. Each image frame is

preprocessed to segment and aligned the collected balloon position data and corresponding catheter behaviour information. The images are structured as multi-channel 2D arrays (e.g., 122×129×2). The input data is processed through two separate convolutional encoding arms, each encoding arm consists of multiple Conv2D layers (filters: 64, 128) with ReLU activation, followed by max-pooling layers (MaxPool) to extract spatial and structural features relevant to catheter navigation. The outputs from both encoding arms are then concatenated. The fully connected layers process this concatenated feature map, utilizing 128 and 64 neurons with ELU activation to capture complex decision-making patterns.

A dropout rate of 0.2 is applied to mitigate overfitting and enhance model generalization. Finally, the output layer employs Softmax activation to classify and predict one of four discrete catheter actions: *Trans*_1: Translation forward, *Trans*_2: Translation backward, *Rota*_1: Rotation right, *Rota*_2: Rotation left. The CNN model is compiled using the Adam optimizer with a learning rate of 1×10^{-3} and the categorical cross-entropy loss function. Training and validation data are converted to NumPy arrays for efficient processing, and a ModelCheckpoint callback is configured to save the best-performing model based on validation accuracy. The model is trained on a dataset consisting of 150 trials across three patient-specific digital phantoms (50 trials per phantom), using a batch size of 32 for 100 epochs. The dataset was randomly split into 80% for training and 20% for validation.

The hardware setup for policy training and image segmentation includes an NVIDIA GeForce RTX 3080 Laptop GPU and a 12th Gen Intel® Core™ i9-12900H CPU. The software environment for the CNN architecture consists of TensorFlow 2.5.0, CUDA Toolkit 11.8, CUDA Deep Neural Network (cuDNN) 8.1.2, and Python 3.9. Fig. 5 presents the training accuracy, validation accuracy and loss of the three models over 100 epochs. The accuracy curves (top) show a steady improvement, with both training and validation stabilizing above 93.5%. The loss curves (bottom) show a consistent decline in both training and validation loss, converging around 0.15. These results indicate that the DL models effectively learned from the input data, exhibiting strong convergence throughout the training process.

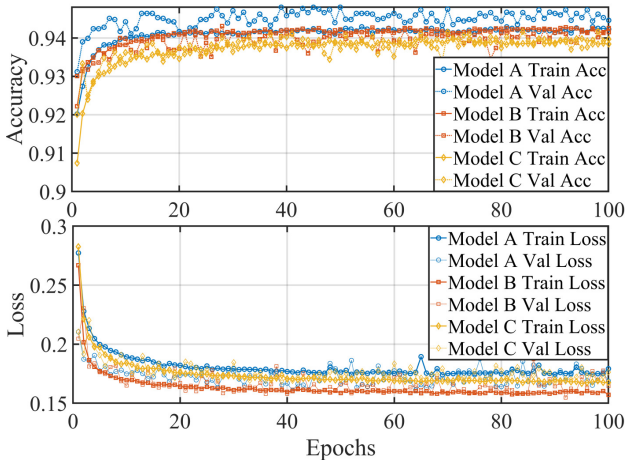
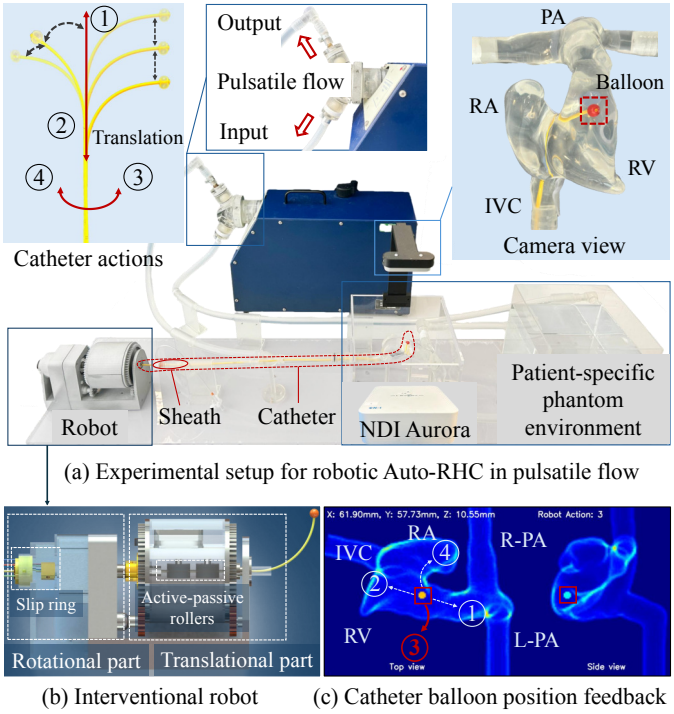


Fig. 5. DL learning curves of the robot prediction models for 3 phantoms. Top figure shows the training and validation accuracy over 100 epochs, and bottom figure shows the variations of training loss and validation loss.



(a) Experimental setup for robotic Auto-RHC in pulsatile flow

(b) Interventional robot (c) Catheter balloon position feedback

Fig. 6. Experiment setup: (a) shows the interventional system that equipped with patient-specific phantom, robot, pulsatile blood pump and tracking system. Communication was established between robot controller, balloon tracking and policy. (b) shows a 2-degree-of-freedom (DoF) interventional robot, including rotational and translational part, that receives real-time commands to manipulate the catheter. (c) is feedback interface of the Auto-RHC with catheter balloon real-time 3D tracking and model-predicted action.

III. Experimental Validation in Patient-Specific Phantoms

This section presents the validation of Auto-RHC in patient-specific phantoms under various conditions. All experiments were conducted in vitro using patient-specific phantoms. Experiment 1: Auto-RHC in static water-filled environment. Experiment 2: Auto-RHC in pulsatile flow phantoms. Experiment 3: Auto-RHC in dynamic pulsatile heartbeat environment and a comparison with expert manual operation.

A. Experiment 1- Auto-RHC in static water-filled phantoms

1) Experimental setup and protocol: Fig. 6 (a) presents the experimental setup for robotic autonomous RHC, including a pulsatile blood pump (Harvard Apparatus, USA), pulmonary artery environment, a 2-DoF Swan-Ganz catheter interventional robot, a camera (Intel Realsense D435i) and the NDI Aurora tracking system. The robot consists of a translational part (active-passive roller set coupled with soft silicone layers) for catheter advancement and retreat movements, and a rotational part integrated with a slip ring for continuous left and right rotation of the catheter. On the software side, the trained model is integrated into a Python

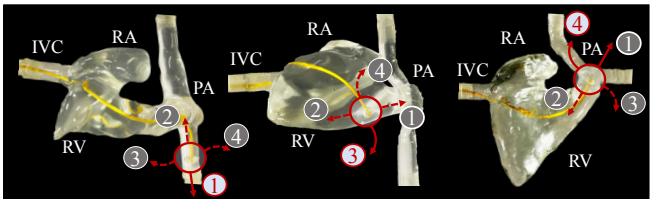
3.9 environment, with communication to the robot controller (Arduino Mega2560) established via Serial communication. As shown in Fig. 7 (top), the patient-specific right heart phantoms were 3D reconstructed from a clinical CT scan and then printed out using VeroClear material with a PolyJet Objet 500 Connex. These three patient-specific phantoms exhibit varying degrees of anatomical complexity in the right ventricle (RV), with Phantom A having the narrowest RV, followed by Phantoms B and C. The constricted RV space poses significant challenges for navigating the flexible pre-curved Swan-Ganz catheter from RV to PA, leading to snagging or buckling (the catheter balloon fails to reach the PA). The phantom's IVC inlet connected the silicone tube (Dimensions: $L \times D = 500\text{mm} \times 16\text{mm}$) that simulate blood vessel. A clinically-used catheter sheath is attached at the other end, enabling the robot to navigate through the vessel and access to IVC.

In Experiment 1, three patient-specific phantoms were filled with water and respectively connected to vessels fitted with catheter sheaths. To enable vision-based real-time 3D tracking of the catheter balloon, a camera was positioned above the patient-specific phantom environment. A deep learning (DL)-based CNN architecture employing the Residual Networks (ResNet) was developed, processing a pre-recorded 2D image dataset of catheter interventions captured by the camera, aligned with 3D balloon position data collected using the NDI Aurora mini tracker (installed in the lumen of catheter tip). The CNN was trained for 800 epochs, with optimal performance observed at 670 epochs. Using Aurora tracker as ground truth, the trained model achieved a localization accuracy within a 3.0 mm error and a mean prediction time of 0.0036 seconds per frame. For further validation, employing the two tracking methods separately, 50 autonomous interventions were performed on each water-filled phantom. Fig. 6(c) presents the feedback interface, providing 3D balloon tracking alongside the predicted robotic actions based on the current position of the catheter balloon. A trial was defined as a failure if snagging on anatomical structures or catheter buckling prevented the robot from advancing to the next configuration, causing the Auto-RHC protocol to stall.

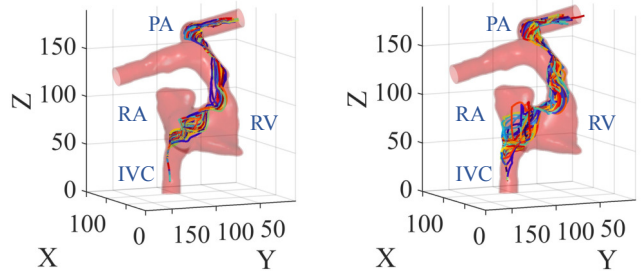
2) Results for Experiment 1: Fig. 7(a) shows the autonomous navigation of the Swan-Ganz catheter in the static water-filled phantom from the IVC to the PA. Table II recorded the results of 50 repetitions of robotic Auto-RHC for each phantom. Both balloon tracking feedback methods showed the success rate of $\geq 96\%$. The successful RHC procedure is defined as the balloon being navigated successfully from the IVC to the target position at the PA. The Aurora tracking demonstrated a 1.67% higher average success rate compared to the vision-based tracking. Notably, both tracking methods achieved a 98% success rate for Phantom B. In terms of Auto-RHC execution time, the vision-based tracking recorded mean time taken of 32.24s, 28.90s, and 26.76s for Phantoms A, B, and C, respectively. Under Aurora tracking, the mean time dropped to 30.70s, 27.42s, and 24.55s, representing an average reduction of 1.74s. In summary, among the 3 patient-specific phantoms under both tracking feedback methods Phantom C demonstrated the shortest execution time and the highest success rate, followed by Phantom B, while Phantom A showed the most time-consuming execution time and a slightly lower success rate compared to Phantoms B and C.

TABLE II
AUTONOMOUS ROBOTIC RHC IN STATIC WATER-FILLED PHANTOMS

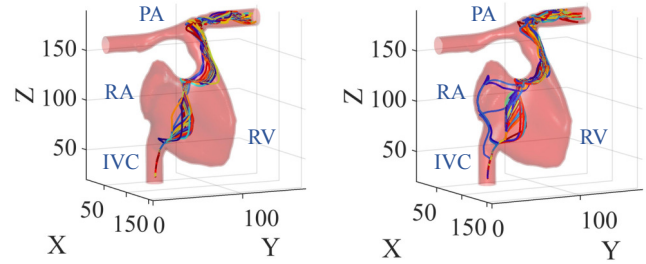
	Water-filled (Camera)			Water-filled (Aurora)		
Cardio phantom	Mean	Min	Succ-rate	Mean	Min	Succ-rate
Patient-specific A	32.24s	30.15s	96%	30.70s	28.69s	98%
Patient-specific B	28.90s	26.55s	98%	27.42s	25.70s	98%
Patient-specific C	26.76s	23.15s	98%	24.55s	21.90s	100%



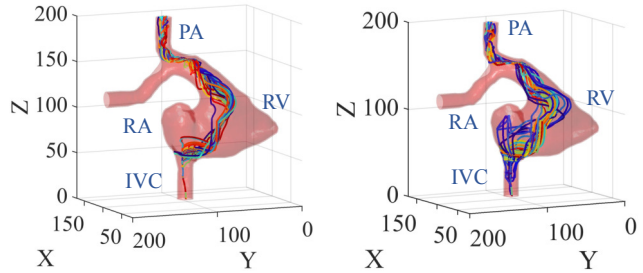
(a) Auto-RHC in static patient-specific phantoms (A-C)



(b) Phantom-A: Robotic Auto-RHC Expert manual operation



(c) Phantom-B: Robotic Auto-RHC Expert manual operation



(d) Phantom-C: Robotic Auto-RHC Expert manual operation

Fig. 7. Experiment 1 and 2 results: (a) Auto-RHC interventions in 3 patient-specific phantoms. (b) Phantom A: Auto-RHC balloon trajectories (left) and manual operation (right). (c) and (d) are the Auto-RHC balloon trajectories (left) and manual operation (right) of Phantom B and C.

TABLE III
AUTONOMOUS ROBOTIC RHC IN PULSATILE FLOW AND DYNAMIC HEARTBEAT PATIENT-SPECIFIC PHANTOMS

Pulsatile Flow	**40 B.P.M**		**50 B.P.M**		**60 B.P.M**		**70 B.P.M**		**80 B.P.M**		**90 B.P.M**	
Cardio Phantom	Mean	Success rate	Mean	Success rate	Mean	Success rate	Mean	Success rate	Mean	Success rate	Mean	Success rate
Patient-specific A	33.40s	96%	33.61s	96%	34.10s	96%	34.42s	96%	35.15s	94%	35.36s	94%
Patient-specific B	30.20s	98%	30.70s	98%	30.82s	98%	31.28s	96%	32.85s	96%	33.28s	96%
Patient-specific C	25.71s	98%	26.86s	98%	27.49s	98%	27.75s	98%	28.08s	96%	28.42s	96%
Dynamic pulsatile heartbeat	35.51s	96%	35.90s	96%	35.75s	96%	36.40s	96%	37.25s	94%	38.72s	94%

B. Experiment 2- Auto-RHC in pulsatile flow phantoms

1) Experimental protocol: A pulsatile blood pump was integrated to generate pulsatile flow within a rigid patient-specific phantom, as illustrated in Fig. 6. The pump's outlet was connected to the phantom's IVC, directing flow through RA, RV and PA. The outlet of left and right pulmonary arteries (L-PA and R-PA) were connected to a fluid tank, which connected back to the blood pump, establishing a closed-loop circulation system. In this experiment, the pump frequency was set from 40 to 90 beats per minute (BPM) in 10 BPM increments. The Auto-RHC procedure was conducted 50 times at each frequency under vision-based balloon real-time tracking feedback. Additionally, a cardiovascular expert with over twenty years of experience was invited to manually operate the catheter for 10 procedures at each frequency for each phantom, recording the balloon navigation data.

2) Results for Experiment 2: Table III presents the results of Experiment 2, summarizing the mean intervention time and success rate of robotic Auto-RHC under pulsatile flow conditions across different frequencies (40–90 BPM). The results indicate that the procedure maintained consistency across all phantoms, with mean time ranging from 33.40 s to 35.36 s for Phantom A, 30.20 s-33.28 s for Phantom B, and 25.71 s-28.42 s for Phantom C. The mean success rates were 95.33% for Phantom A, 97.0% for Phantom B, and 97.67% for Phantom C. Consistent with observations in the static water-filled environment, Phantom A showed the most time-consuming, with an average time of 34.34 s, followed by Phantom B at 31.52 s, while Phantom C demonstrated the shortest intervention time at 27.39 s.

Fig. 7 (b)-(d) present of the balloon trajectories of robotic Auto-RHC and expert manual operation for patient-specific Phantom A-C in pulsatile flow environments, 10 trajectories were randomly selected for each pulsating frequency, totaling 60 trajectories. Root Mean Square Deviation (RMSD) is employed to evaluate consistency of all trajectory datasets, which are resampled to a uniform length $K = 200$ sampling point. For trajectory sampling point in k (with $K = 1, 2, \ldots, K$), the mean trajectory across all N robotic Auto-RHC trajectories is calculated as:

$$\{\overline{X}(k), \overline{Y}(k), \overline{Z}(k)\} = \frac{1}{N}\sum_{i=1}^{N}\{X_i(k), Y_i(k), Z_i(k)\} \tag{9}$$

Where $\overline{X}(k), \overline{Y}(k), \overline{Z}(k)$ are the mean X, Y, and Z coordinates at k, $X_i(k), Y_i(k), Z_i(k)$ are the i-th trajectory, N is the total number of trajectories.

The overall RMSD across all trajectories is calculated as:

$$\text{RMSD} = \left(\frac{1}{N \times K}\sum_{i=1}^{N}\sum_{k=1}^{K}\left[\left(X_i(k) - \overline{X}(k)\right)^2 + \left(Y_i(k) - \overline{Y}(k)\right)^2 + \left(Z_i(k) - \overline{Z}(k)\right)^2\right]\right)^{1/2} \tag{10}$$

The results indicate that for patient-specific phantom A, the RMSD of robotic Auto-RHC is 6.812mm, compared to 10.420mm for expert manual operation. For Phantom B, Auto-RHC achieves an RMSD of 6.525mm, while manual RHC results in 10.772mm. In Phantom C, the RMSD for Auto-RHC is 7.490mm, whereas the manual procedure yields 12.265mm. These findings demonstrate a notable improvement in trajectory consistency with Auto-RHC, achieving RMSD reductions of 34.63% for Phantom A, 39.43% for Phantom B, and 38.93% for Phantom C compared to the clinical expert manual operation.

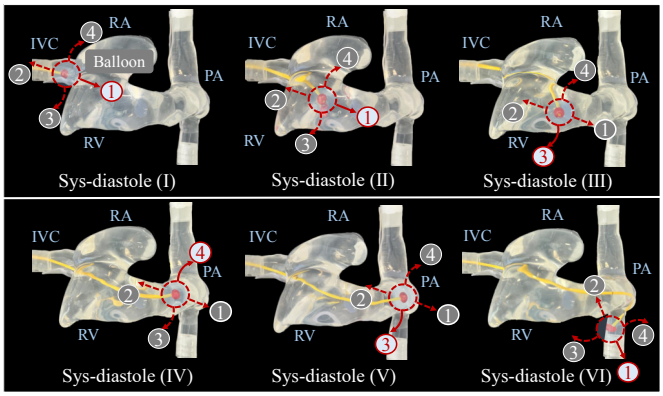
C. Experiment 3- Auto-RHC in dynamic heartbeat phantom

1) Experimental protocol: To realistically simulate the dynamic physiological environment of the patient heart (heart rates: 40-90 BPM, systole/diastole range: 50/50-25/75), a soft, transparent silicone phantom of the patient-specific phantom A was 3D printed by specialized commercial manufacturer (Trandomed 3D Inc., China) [19]. As shown in Fig. 8(a), the silicone material was selected based on its mechanical properties, including a Shore hardness of 43, tensile strength of 6.17 MPa, and tensile elongation of 294% the thickness of the printed heart model was set to 1.5 mm. Phantom A was selected for validation in this experiment due to its anatomically complex structure, including a narrowed RV and PA stenosis, which posed greater challenges for catheter navigation compared to Phantoms B and C, as demonstrated in Experiments 1 and 2. The soft patient-specific phantom was connected to a pulsatile blood pump, generating a realistic dynamic pulsating heartbeat environment (See the supplementary video) and involves both systole and diastole (Sys-diastole) phases. The heartbeat frequency was systematically varied from 40 to 90 BPM in 10 BPM increments, representing a range of patient-related heart rates. At each frequency, 50 robotic Auto-RHC procedures were performed. During the catheter intervention, key parameters, including balloon position, intervention duration were recorded to evaluate the Auto-RHC performance under dynamic heartbeat conditions. For comparison, the expert's manual operation was conducted

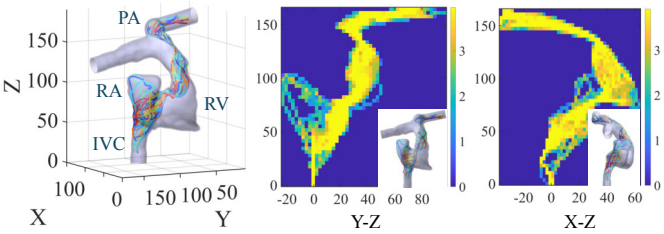
10 operation at each heartbeat rate, with positional data for catheter balloon navigation recorded for all trials.

2) Results for Experiment 3: Tab. III (bottom section) presents the results of Experiment 3, demonstrating that the Auto-RHC success rate in dynamic heartbeat scenarios remains between 94% and 96% across various heartbeat frequencies, with a mean success rate of 95.33%. As the heartbeat frequency increases, the success rate gradually declines from 96% to 94%. Additionally, the execution time progressively increases with higher heartbeat frequencies, ranging from 35.51 s at 40 BPM to 38.72 s at 90 BPM, with a total variation of 3.5 s. The average time taken across heartbeat frequencies from 40 BPM to 90 is 36.58 s.

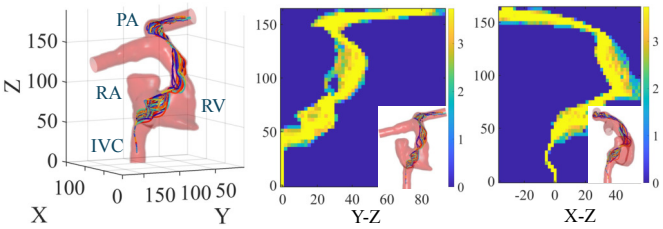
Fig. 8 (a) illustrates the robotic autonomous catheter navigation within the patient-specific dynamic heartbeat phantom. Guided by vision-based balloon tracking feedback, the trained decision-making policy delivers real-time action commands to the robot, enabling catheter navigation from IVC to PA by translational and rotational movements. Fig. 8(b) presents the catheter balloon trajectories from expert manual interventions, Fig. 8(c) shows 60 Auto-RHC trajectories, with 10 randomly selected for each heartbeat frequency. The results indicate that the RMSD for expert manual interventions is 12.720 mm, whereas robotic Auto-RHC achieves an RMSD of 7.426 mm for all data, demonstrating a 41.61% improvement in trajectory consistency under dynamic heartbeat conditions.



(a) Autonomous robotic RHC in dynamic heartbeat phantom



(b) Expert manually operated balloon trajectories and heat map



(c) Robotic Auto-RHC balloon trajectories and heat map

Fig. 8. Experiment 3 results: (a) Autonomous robotic RHC in dynamic heartbeat patient-specific phantom. (b) Expert manually operated balloon trajectories. (c) Robotic Auto-RHC balloon trajectories and heat map.

D. Discussion

In Experiment 1, employing the NDI Aurora sensor for real-time feedback resulted in slightly shorter intervention times (by 2.28s) and a 1.3% increase in success rates compared to the vision-based feedback method across all patient-specific phantoms. This difference is primarily attributed to the higher tracking accuracy provided by the Aurora sensor, which enabled the policy network to predict more precise robotic actions based on the real-time position of the catheter balloon. Embedding the mini NDI Aurora tracker within the lumen of the catheter tip is technically feasible. It is important to note, however, that Swan-Ganz catheters typically comprise two lumens: one for balloon inflation/deflation and the other for drug administration or saline injection. Using one lumen for tracker integration (along with the tracker's wired connection to NDI receiver) would compromise the catheter's clinical functionality. External tracker integration might affect the catheter's dimensional conformity, obstructing sheath-based delivery. Hence, we employed an image- and DL-based balloon tracking framework in vitro study, offering a non-invasive and clinically transferable solution [20]. Nevertheless, transferring this approach to clinical imaging modalities, such as X-ray fluoroscopy, can introduce technical challenges including latency, limited image resolution/dimension, and real-time computational demands. Additionally, phantom geometry significantly influenced both the success rates and intervention times of the Auto-RHC procedures. Among the three phantoms, Phantom C exhibited the highest success rate (98%) and the shortest average execution time, followed by Phantoms B and A. The reduced performance observed in Phantom A is attributed to its anatomical features, specifically the narrower RV and PA. These geometric constraints limited the maneuverability of the pre-curved Swan-Ganz catheter, requiring continuous rotational adjustments to successfully traverse RV and reach PA. This additional manipulation introduced greater procedural complexity, increasing both the likelihood of failure and the execution time.

In Experiment 2, compared to Experiment 1, the average success rate of autonomous Auto-RHC in a rigid patient-specific phantom environment with pulsatile water flow showed a slight decline across all three phantoms (from 96% to 95.33% for Phantom A, 98% to 97% for Phantom B, and 98% to 97.67% for Phantom C). Additionally, the execution time increased for all phantoms. This performance drop is primarily attributed to the dynamic nature of the pulsatile flow, which caused the flexible Swan-Ganz catheter to wobble in synchrony with the pulse cycles. Such oscillations disrupted the real-time vision-based balloon tracking, making it more challenging for the trained catheter policy to accurately interpret the catheter's position. The impact of these disturbances became more pronounced with increasing pulse frequencies, as higher-frequency pulsations led to greater jittering of the

catheter balloon. This, in turn, interfered with the policy's real-time decision-making process, resulting in a gradual decline in the overall success rate and extended execution times. Furthermore, RMSD analysis comparing the expert's manual trajectories with the robotic Auto-RHC balloon trajectories reveals a significant improvement in catheter navigation consistency. This enhanced precision reduces procedural variability, thereby minimizing the risk of perforation and improving the overall safety of Auto-RHC compared to manual RHC. Regarding RMSD variation, Phantom C exhibited a higher RMSD compared to Phantoms A and B. This trade-off suggests that, without compromising procedural success, the robot decision-making model prioritized efficient task completion (navigating the balloon to the PA) over fine trajectory optimization.

In Experiment 3, the success rate of Auto-RHC in a dynamic, pulsatile heartbeat patient-specific environment exhibited a slight decrease, from 96% in the static scenario to 94%. This reduction is primarily attributed to the challenges of vision-based 3D real-time balloon tracking in a dynamic cardiac environment. Specifically, the continuous systole and diastole of the right heart cause periodic oscillations of the catheter's curved tip and balloon, influencing the accuracy of balloon tracking. These dynamic deviations interfere with the real-time feedback essential for the catheter decision-making policy, leading to reduced tracking accuracy and success rate. The tracking challenge becomes more pronounced as pulse frequencies increase, to address this challenge, the development of a dynamic phantom-based balloon tracking model is required. Despite these challenges, the proposed SOFA-based framework, integrated with CNN, achieved a mean intervention success rate of 95.33% in dynamic environment. Fig. 8 (c) show the consistency improved by 41.61% compared to manual operation, maintaining the Auto-RHC reliable performance under pulsatile heartbeat environment. While the current patient-specific phantom realistically simulates heartbeats, it omits the tricuspid valve. This limitation reduces anatomical complexity and may affect the evaluation of robotic Auto-RHC navigation performance.

IV. Conclusion and Future Work

This paper presents the transfer from patient-specific right heart digital twins to real-world robotic Auto-RHC, enabling the autonomous navigation of the Swan-Ganz catheter in both static and dynamic heartbeat patient-specific phantoms. By accurately modelling of the Swan-Ganz catheter and creating a RHC digital twin in the SOFA framework, the physical RHC procedure can be precisely replicated in the virtual environment. Operators can collect positional and behavioral datasets of catheter–phantom interactions by performing interventions in the digital twin environment. This dataset served as the input for the CNN architecture, which was used to train a robotic decision-making policy. In the real-world robotic system, a real-time communication framework was established between the robot controller, the trained prediction model and the balloon tracking system, enabling the robot to receive commands and autonomously navigate the catheter from the IVC to the PA. Experimental validation demonstrated that Auto-RHC achieved high success rates in patient-specific phantoms, with $\geq 96\%$ in static environments and $\geq 94\%$ in dynamic pulsating heartbeat conditions. Additionally, Auto-RHC exhibited a significant improvement ($\geq 34.63\%$) in trajectory consistency compared to expert manual operations, a key factor in reducing procedural variability and enhancing the safety of clinical right heart catheterization.

Future work will focus on optimizing patient-specific in vitro right heart models by incorporating region-specific mechanical properties and key physiological features, such as the tricuspid valve, delivering a high-fidelity phantom environment for Auto-RHC. In addition, a robotic operating console will be developed to introduce a clinician-in-the-loop safety mechanism. This shared autonomy framework will allow clinicians to intervene when the autonomous navigation system encounters difficulties, enabling a safer, more efficient, and clinically accessible robotic right heart catheterization procedure.

V. Acknowledgment

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising. The authors would like to thank the Royal Free Hospital (London, UK) for providing patient-specific cardiac models. The authors also gratefully acknowledge Prof. Vivek Muthurangu for the access to the pulsatile blood pump and technical support for clinical RHC workflows.

References

[1] S. Martin, A. Aday, Z. Almarzooq, C. Anderson, and P. Arora, "2024 heart disease and stroke statistics: A report of us and global data from the american heart association," *Circulation*, vol. 149, no. 8, pp. e347–e913, 2024.

[2] G. Mensah, V. Fuster, C. Murray, G. Roth, Y. Abate, M. Abbasian, F. Allah, A. Abdollahi, and M. Abdollahi, "Global burden of cardiovascular diseases and risks, 1990-2022," *Journal of the American College of Cardiology*, vol. 82, no. 25, pp. 2350–2473, 2023.

[3] M. Omisore, O. 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.

[4] Y. Cho, J.-H. Park, J. Choi, and D. E. Chang, "Sim-to-real transfer of image-based autonomous guidewire navigation trained by deep deterministic policy gradient with behavior cloning for fast learning," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 3468–3475.

[5] L. Karstensen, J. Ritter, J. Hatzl, T. Pätz, J. Langejürgen, C. Uhl, and F. Mathis-Ullrich, "Learning-based autonomous vascular guidewire navigation without human demonstration in the venous system of a porcine liver," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 11, pp. 2033–2040, 2022.

[6] Y. Wang, J. Wang, Y. Li, T. Yang, and C. Ren, "The deep reinforcement learning-based vr training system with haptic guidance for catheterization skill transfer," *IEEE Sensors Journal*, vol. 22, no. 23, pp. 23 356–23 366, 2022.

[7] Z. Mei, J. Wei, S. Pan, H. Wang, D. Wu, Y. Zhao, G. Liu, and S. Guo, "Transferring virtual surgical skills to reality: Ai agents mastering surgical decision-making in vascular interventional robotics," *IEEE/ASME Transactions on Mechatronics*, pp. 1–12, 2024.

[8] V. Scarponi, M. Duprez, F. Nageotte, and S. Cotin, "A zero-shot reinforcement learning strategy for autonomous guidewire navigation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, pp. 1185–1192, 2024.

[9] H. Robertshaw, L. Karstensen, B. Jackson, A. Granados, and T. C. Booth, "Autonomous navigation of catheters and guidewires in mechanical thrombectomy using inverse reinforcement learning," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, no. 8, pp. 1569–1578, 2024.

[10] T. Jianu, B. Huang, T. Vo, M. Vu, J. Kang, H. Nguyen, O. Omisore, P. Berthet-Rayne, S. Fichera, and A. Nguyen, "Autonomous catheterization with open-source simulator and expert trajectory," 2024, *arXiv:2401.09059*.

[11] H. Su, K.-W. Kwok, K. Cleary, I. Iordachita, C. Cavusoglu, J. P. Desai, and G. S. Fischer, "State of the art and future opportunities in mri-guided robot-assisted surgery and interventions," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 968–992, 2022.

[12] A. Pore, Z. Li, D. Dall'Alba, A. Hernansanz, E. De Momi, A. Menciassi, A. Casals Gelpí, J. Dankelman, P. Fiorini, and E. V. Poorten, "Autonomous navigation for robot-assisted intraluminal and endovascular procedures: A systematic review," *IEEE Transactions on Robotics*, vol. 39, no. 4, pp. 2529–2548, 2023.

[13] W. Chi, G. Dagnino, T. Kwok, A. Nguyen, D. Kundrat, K. Abdelaziz, Mohamed, 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)*, 2020, pp. 2414–2420.

[14] H. Rafii-Tari, J. Liu, S.-L. Lee, C. Bicknell, and G.-Z. Yang, "Learning-based modeling of endovascular navigation for collaborative robotic catheterization," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, 2013, pp. 369–377.

[15] W. Chi, J. Liu, H. Rafii-Tari, C. Riga, C. Bicknell, and G.-Z. Yang, "Learning-based endovascular navigation through the use of non-rigid registration for collaborative robotic catheterization," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 04, pp. 855–864, 2018.

[16] J. Lai, T.-A. Ren, W. Yue, S. Su, J. Chan, and H. Ren, "Sim-to-real transfer of soft robotic navigation strategies that learns from the virtual eye-in-hand vision," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 2, pp. 2365–2377, 2024.

[17] F. Faure, C. Duriez, H. Delingette, J. Allard, B. Gilles, S. Marchesseau, H. Talbot, H. Courtecuisse, G. Bousquet, I. Peterlík, and S. Cotin, "Sofa, a multi-model framework for interactive physical simulation," 2011.

[18] C. Neumann, M. Breil, A. Schild, A. Schenk, P. Jakobs, M. Mikus, and E. Schindler, "Central venous catheter tip positioning using ultrasound in pediatric patients: A prospective observational study," *Pediatric Anesthesia*, vol. 34, no. 4, p. e14864, 2024.

[19] X. Zhang, Y. Che, C. Li, X. Cao, X. Wang, Y. Zhang, G. Li, J. Zhu, M. Luo, and C. Shu, "Prediction of iliac limb occlusion after endovascular aneurysm repair for abdominal aortic aneurysm by anatomical and near-wall hemodynamic characteristics combining numerical simulation and in vitro experiment," *Computer Methods and Programs in Biomedicine*, vol. 268, p. 108845, 2025.

[20] I. Vernikouskaya, D. Bertsche, T. Dahme, and V. Rasche, "Cryo-balloon catheter localization in x-ray fluoroscopy using u-net," *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 8, pp. 1255–1262, 2021.