# Reinforcement learning for navigation in percutaneous coronary arteries interventions

Serena Elzein, Luc Duong
Dept. of Software and IT Engineering, École de Technologie Supérieure, Montréal, Canada

## ABSTRACT

Percutaneous coronary intervention (PCI) is a widely used minimally invasive procedure to treat coronary artery disease. Despite its benefits, PCI relies heavily on the dexterity of the operator and can be highly challenging during navigation. Advances in robot-assisted PCI have shown promise in reducing these risks and enhancing procedural precision. This study presents the development and in silico evaluation of a simulation environment equipped with reinforcement learning (RL) to enable autonomous catheter navigation for PCI, with potential applications in training operators and assisting in real-time procedure guidance. Developed using Unity 3D and Unity ML-Agents, our model utilizes a simulation environment to train RL agents for catheter guidance. The navigation within the coronary arteries was modeled using Unity's game engine, which allows for realistic catheter movements and collision detection with vessel walls using mesh colliders. The automatic catheterization employs a goal-based binary function, reinforced by a checkpoint reward system that directs the agent's movements toward successful navigation. To evaluate the model, extensive simulation trials were conducted with different movement boundaries to track learning progress and refine training strategies. The results from these in silico trials suggest significant improvements in procedural safety and efficiency, indicated by a reduction in navigation error from an initial average of 0.05 (±0.01) mm to less than 0.01 (±0.002) mm. Cumulative rewards steadily increased, showing a final average of 0.5 (±0.1) mm in reward values. These metrics demonstrate the model's ability to adapt and optimize its performance for a range of catheter navigation scenarios. However, the evaluation is limited to a virtual environment, and further work is necessary to assess how these findings translate to real-world clinical applications. In particular, integrating this RL-based approach with live PCI procedures will require addressing patient-specific variability, real-time physiological changes, and ensuring safe interaction between the AI agent and human operators during procedures. This study represents an important step toward incorporating AI into cardiac healthcare, but practical implementation in clinical settings will require further investigation, including experimental or clinical validation. Future research should focus on testing the method alongside human operators in controlled clinical environments to evaluate its effectiveness as a real-time guidance tool during PCI.

Keywords: Reinforcement learning, Navigation guidance, Coronary artery, Image-guided interventions

## 1. DESCRIPTION OF PURPOSE

Percutaneous coronary intervention (PCI) is a minimally invasive procedure to treat coronary artery disease by restoring blood flow to the heart. Despite its benefits, PCI poses risks like radiation exposure to medical staff and patients, and complications such as bleeding and stroke. Conventional PCI uses X-ray fluoroscopy, exposing patients and medical personnel to potentially harmful radiation. Recent advancements in robot-assisted PCI aim to enhance procedural precision and reduce these risks. Robotic systems, using technologies like electromyography and accelerometers, improve control and accuracy in guidewire manipulation. Analysis of interventionalists' natural behaviors has shown that recognizing motion patterns of endovascular tools is crucial for developing precise robotic systems [2]. Reviews of current and emerging robot-assisted endovascular catheterization technologies highlight significant improvements in procedural outcomes through enhanced precision and control [3]. The third-generation magnetic navigation system significantly reduces procedure duration in remote magnetic-guided catheter ablation of atrial fibrillation, showcasing the efficiency gains possible with advanced robotic systems [4]. Additionally, predictive filtering in motion compensation with steerable cardiac catheters can further enhance the precision of robotic-assisted procedures [5]. In the context of PCI, integrating artificial intelligence (AI), particularly reinforcement learning (RL), presents a promising avenue for enhancing precision and safety. RL trains an agent to make a sequence of decisions by rewarding desirable actions and penalizing undesirable ones. This real-time adaptation and optimization capability can significantly enhance catheterization precision, reduce procedural risks, and improve patient outcomes.

Medical Imaging 2025: Image-Guided Procedures, Robotic Interventions, and Modeling,
edited by Maryam E. Rettmann, Jeffrey H. Siewerdsen, Proc. of SPIE Vol. 13408,
134082P · © 2025 SPIE · 1605-7422 · doi: 10.1117/12.3047034

The goal of this study is to develop a RL-based model to optimize catheterization during PCI. Using Unity 3D and the Unity ML-Agents library, we have created a high-fidelity simulation environment to train the RL model. The manual catheterization component provides a baseline for developing automatic techniques, leveraging both human expertise and AI optimization. Our goal is to demonstrate RL's potential to improve the reliability, safety, and efficiency of PCI procedures, setting new standards in cardiac interventions. By addressing current PCI limitations and harnessing advanced AI techniques, this research aims to enhance cardiac care, ensuring better patient outcomes and advancing technology integration in medical procedures. This work is organized as follows: Section 2 presents the simulation environment setup, Section 3 reports the results obtained on coronary artery phantom. Section 4 discusses the results and concludes by outlining the novelty of our approach and future works.

## 2. METHODS

### 2.1 Design of a manual catheterization 3D environment

To enhance navigation guidance during PCI, we designed a 3D environment using a 3D game engine (Unity Technologies, San Francisco, USA). A 3D model was created by segmenting a CT acquisition of a physical phantom of the coronary arteries (Fig. 1). Navigation pathways inside the coronary arteries were obtained using centerline extraction of the lumen of the coronary artery with the Vascular Modeling Toolkit (VMTK). The manual catheterization component utilizes Unity's Standard 3D Package, enhanced with specialized assets to create a high-fidelity simulation environment. We simulated the movements of the guidewire and interactions essential for catheterization. This includes the hand movement for manipulating the guidewire, providing users with a hands-on feel for the procedure. Additionally, collision detection with mesh colliders replicates the tactile feedback experienced during the procedure. These effects are essential for creating a realistic virtual environment and for allowing proper movement of the catheter within safe bounds in the artery (Fig. 1). The catheterization procedure is controlled using user input, allowing medical students, trainees, and professionals to virtually perform catheterization and gain a realistic sense of navigating within an artery without causing damage. This immersive experience aids in understanding how to properly navigate towards stenosis. Furthermore, we implemented both external and internal points of view (POVs) to allow for comprehensive visualization of the entire heart model (Fig. 2).

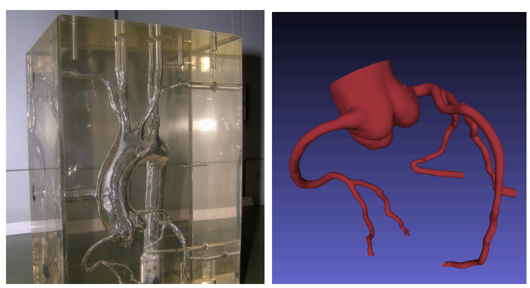

*Figure 1: Design of the 3D environment for navigation guidance for PCI. A physical phantom (left) is filled with contrast agent and acquired using CT imaging. A 3D model of the coronary artery is segmented from the CT acquisition (right). The centerlines of the 3D model is used for defining the navigation path of the guide wire inside the coronary arteries.*

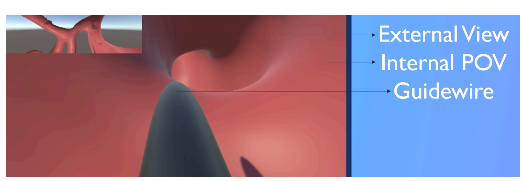


*Figure 2: External and internal points of view of catheterization*

### 2.2 Automatic navigation using reinforcement learning

The automatic catheterization module progresses from manual navigation, transitioning to an environment where a machine learning model, trained via reinforcement learning, autonomously navigates. The model follows a policy dictated by a goal-based binary function, aiming to replicate the precision of manual catheterization without human input.

$$R(s,a)=\begin{cases}1 & \text{if goal achieved (no collision and passing through checkpoint)}\\ 0 & \text{if action neutral}\\ -1 & \text{if action leads to collision}\end{cases}$$

Where *R* is the reward function, *s* represents the state, and *a* represents the action. A reward of 1 is given if the guidewire passes through a checkpoint without hitting boundaries, and a reward of -1 is given if a collision occurs. These rewards are based on distance computations between the guidewire position and the vessel wall normals.

A nuanced checkpoint reward system (Fig. 3) reinforces successful navigation strategies, honed through extensive training sessions managed via command line. By setting specific goals or 'checkpoints' within the environment, the model receives feedback when it achieves certain milestones, reinforcing the desired behavior. This structured reward system accelerates the learning process, as the model iteratively refines its strategy to maximize cumulative rewards. The system rewards the model for successfully navigating the heart's complex anatomy, reinforcing the precise movements required for PCI procedures. When the guidewire moves through the vessel, the Rigidbody component detects collisions by monitoring changes in the guidewire's position. If a collision occurs, the guidewire's position is reset to the previous valid position. This process calculates the exact distance traveled before hitting the boundary, helping the agent understand spatial constraints and adjust its actions. By computing the distance between the guidewire and vessel wall normals, the agent optimizes its navigation strategy to avoid collisions. This feedback loop enhances learning by providing precise spatial information for refining movements.

In the implementation on heart models, we designed a 3D scene setup with standardized elements to simulate the environment, alongside an agent setup embodying the guidewire for catheterization. The environment's boundaries were defined to facilitate precise agent movement. We implemented collision detection in Unity by adding Mesh Colliders to the boundaries of the 3D scene, which detected when the guidewire touched the vessel walls. The simulation environment represented a range of coronary artery structures, with the inclusion of multiple clinical cases to reflect diverse anatomical challenges. Manual boundaries were created on the inside of the mesh by inverting the normals, effectively flipping the mesh inside out so that the inverted normals could serve as boundaries. This setup enabled precise collision detection and provided immediate feedback to the agent, aiding in the training process. These environments were replicated to quicken the training process, enabling the reinforcement learning model to rapidly assimilate and apply knowledge across diverse scenarios, ultimately streamlining catheterization training efficacy. The approach was implemented using Unity's ML-Agents Library with PyTorch and TensorFlow for the model's learning architecture. These frameworks were employed to monitor the agent's training progress, recording successful navigations without collisions. They evaluated the effectiveness of the goal-based binary function policy, tracked the agent's cumulative rewards, and were instrumental in fine-tuning the training regimens to enhance the agent's performance.

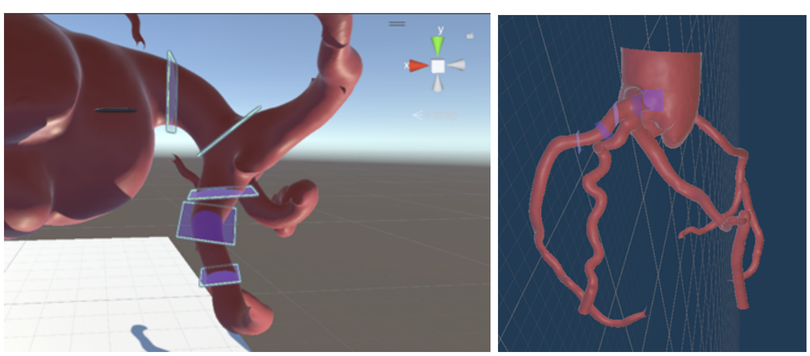

*Figure 3: Checkpoint reward system for the agent using goal based binary function in a 3D model of the coronary artery model.*

# 3. RESULTS

## 3.1 Evaluation of the RL model

The RL agent's performance was evaluated through policy and value losses, and cumulative rewards. The policy and value losses showed a decrease from initial values of 0.05 (±0.01) mm to less than 0.01 (±0.002) mm, suggesting a significant improvement in the model's accuracy. This reduction in loss indicates the agent's increased precision in navigating coronary arteries. Cumulative rewards increased steadily, with final trials showing an average reward increase of 0.5 (±0.1) mm, reflecting the agent's ability to perform more efficient, successful catheterization steps. These improvements indicate that the model is making fewer errors and optimizing its navigation strategy. The fluctuations in policy and value losses during different training phases suggest the agent's evolving strategy to minimize errors. As shown in Figure 4, both policy and value losses decline over time, reinforcing the model's improving performance with more training. The upward trend in cumulative rewards indicates that the agent is increasingly achieving its goals and navigating within the defined boundaries of the heart model environment. These metrics demonstrate the model's improving proficiency in autonomous navigation for PCI procedures. Overall, the RL model shows considerable promise in improving the safety and accuracy of PCI, potentially reducing operator risks and increasing procedural efficiency in clinical practice.

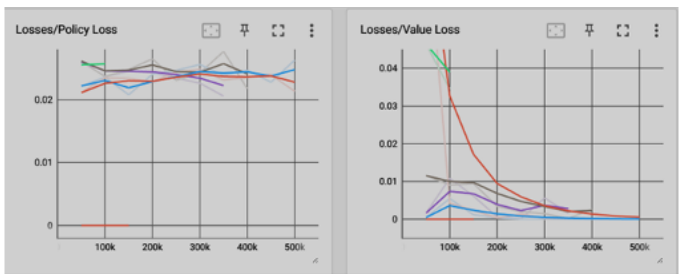


*Figure 4: Variation of Policy and Value Losses of Agent over time for different trainings*

The cumulative reward trends upwards, suggesting that the agent is increasingly achieving its goals within the training parameters set in the heart model environments. This indicates that the agent is getting better rewards as it takes more steps in the environment. These metrics collectively show the agent's improving proficiency in autonomous navigation for catheterization. This suggests that incorporating RL into PCI can lead to better procedural outcomes, reduced risks, and enhanced efficiency in clinical practice.

# 4. NEW WORK TO BE PRESENTED

This study presents a new approach for using an RL-based agent for navigation guidance in coronary arteries. Recent advancements in reinforcement learning (RL) have demonstrated significant potential in enhancing autonomous guidewire navigation in medical procedures. A zero-shot RL strategy allows agents to navigate guidewires without prior training on specific anatomical structures by leveraging a shape-invariant observation space and a sophisticated reward function [1]. Similarly, VesNet-RL has proven the effectiveness of simulation-based RL for real-world ultrasound probe navigation, showing its suitability for medical tasks [8]. Building on these principles, our study integrates advanced RL techniques to optimize catheterization in PCI. We designed a 3D environment for simulating manual navigation and implemented collision detection to ensure the guidewire avoids vessel walls. While the RL model is trained and tested entirely within a virtual environment, it is crucial to address how these virtual outcomes will translate to real-world clinical settings. Currently, the clinical translation may be constrained by factors such as patient-specific variability, the complexity of handling real-time physiological changes, and the inherent limitations of virtual simulations that may not capture all the nuances of human anatomy and operator variability. Our work represents one of the first integrations of a game engine environment and RL for PCI navigation guidance, aiming to set new standards in cardiac care, but further validation in real clinical scenarios will be necessary to assess the robustness and reliability of this approach in patient care.

## 5. CONCLUSION

This study shows that the reinforcement learning model improves the accuracy and safety of catheterization in PCI procedures. Using a 3D simulation environment and a checkpoint reward system, we trained an RL agent to navigate catheterization paths with precision and reduced risks. Our findings align with advancements in motion pattern recognition and emerging technologies for better procedural outcomes. Studies have shown that recognizing motion patterns of endovascular tools and utilizing emerging technologies can significantly improve procedural outcomes [2,3]. Additionally, efficiency gains observed in remote magnetic-guided catheter ablation and precision enhancements from predictive filtering support the benefits of integrating RL into PCI. Specifically, remote magnetic-guided catheter ablation has demonstrated significant reductions in procedure duration, and predictive filtering has enhanced motion compensation in steerable cardiac catheters [4,5]. While the results from simulation are promising, it is important to highlight the gap between virtual training and actual clinical application. Translating this model into real patient settings will require further clinical trials, considering factors like real-time adjustments to dynamic patient conditions (e.g., cardiorespiratory motion) and handling unexpected anatomical variations. This study indicates that RL can be effectively applied to enhance complex medical procedures, providing real-time adaptation and optimization. Future work will focus on integrating adaptive RL techniques, investigating the impact of cardiorespiratory motion, and applying this approach to other minimally invasive procedures to further improve clinical outcomes, with the aim of bridging the gap between simulation and actual clinical practice.

## REFERENCES


1. V. Scarponi, M. Duprez, F. Nageotte, S. Cotin, A zero-shot reinforcement learning strategy for Autonomous Guidewire Navigation. arXiv.org (2024).
2. Zhou, X. H., Xu, X., Chen, J., & Zhang, Y. (2019). Analysis of interventionalists' natural behaviors for recognizing motion patterns of endovascular tools during percutaneous coronary interventions. IEEE Transactions on Biomedical Circuits and Systems, 13(2), 330-342.
3. Rafii-Tari, H., Payne, C. J., & Yang, G. Z. (2014). Current and emerging robot-assisted endovascular catheterization technologies: A review. Annals of Biomedical Engineering, 42(4), 697-715.
4. Maurer, T., Kuck, K. H., & Hoffmann, B. A. (2017). Significant reduction in procedure duration in remote magnetic-guided catheter ablation of atrial fibrillation using the third-generation magnetic navigation system. Journal of Interventional Cardiac Electrophysiology, 49(3), 219-226.
5. Loschak, P. M., Degirmenci, A., & Howe, R. D. (2017). Predictive filtering in motion compensation with steerable cardiac catheters. 2017 IEEE International Conference on Robotics and Automation (ICRA) (pp. 4830-4836). IEEE.
6. Y. R. Manda, K. M. Baradh, Cardiac catheterization risks and complications (2023), (available at https://www.ncbi.nlm.nih.gov/books/NBK531461/).
7. Y. Bi et al., VesNet-RL: Simulation-based Reinforcement Learning for real-world US probe navigation. IEEE Robotics and Automation Letters
8. F. Azizmohammadi, I. Navarro Castellanos, J. Miró, P. Segars, E. Samei, L. Duong. Patient-specific cardio-respiratory motion prediction in X-ray angiography using LSTM networks. Phys Med Biol. 2023 Jan 5;68(2):10.1088/1361-6560/acaba8