POLITECNICO
MILANO 1863

Path Planning via Reinforcement Learning for Robotic Catheters Performing Percutaneous Coronary Interventions in a Deformable Environment

Advisor: Prof. Elena De Momi
Co-advisor: Zhen Li, Ph.D. candidate

Author: Chiara Lambranzi 944336
Academic Year 2021/22

Department of Electronics, Information and Bioengineering
Master of Science – Biomedical Engineering

Introduction

Clinical scenario

- Cardiovascular diseases main cause of death in developed countries[1]
- 147.438 PCIs in Italy in 2021[2]

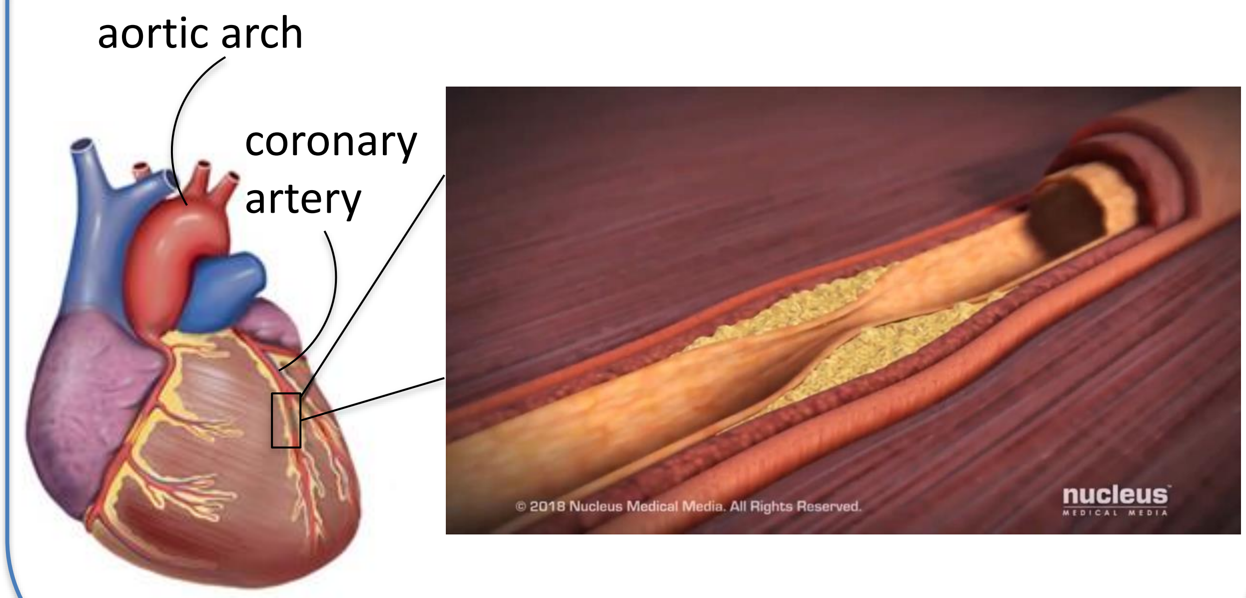
Risks

- X-ray and constrast dye exposure
- Collision with delicate anatomy

Solution

- Aortic model from MRI
- Guidance with path planning in PCI procedure to reduce collision

PERCUTANEOUS CORONARY INTERVENTION (PCI)

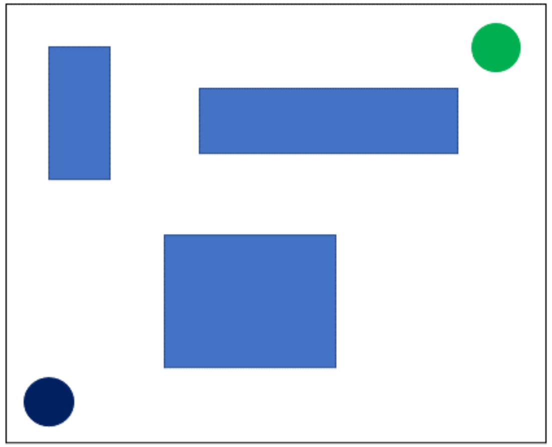


[1] OECD (2021)
[2] GICI (2022)

State of the art

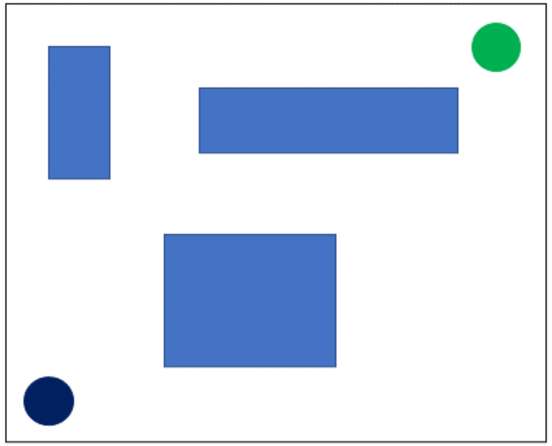
Path planning: find a collision-free motion between an initial and a final configuration within a specified environment

Node based



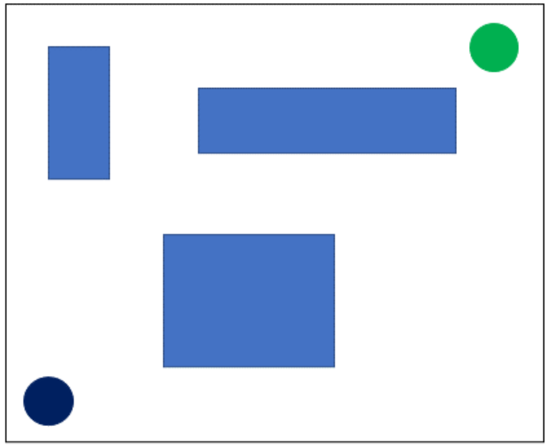
Dijkstra's Algorithm

Sampling based



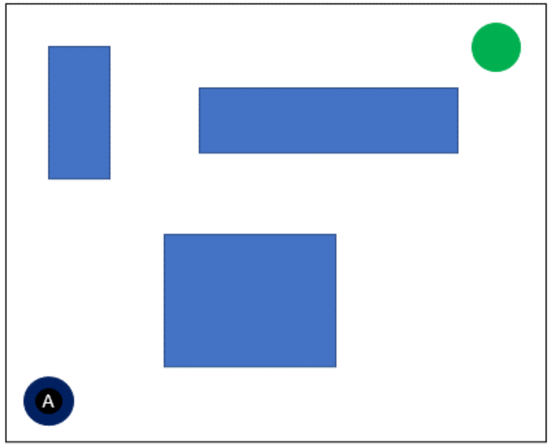
Rapidly-exploring Random Tree

Optimization based



Bézier interpolation

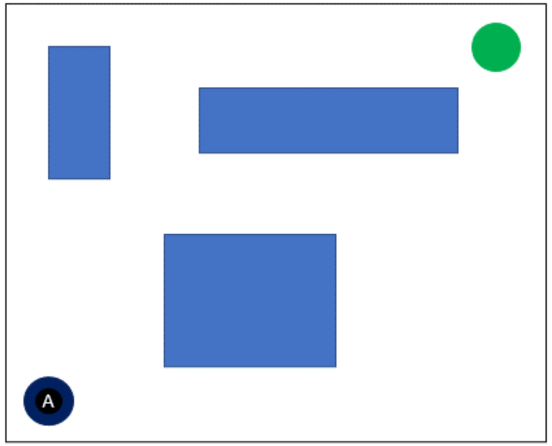
Learning based



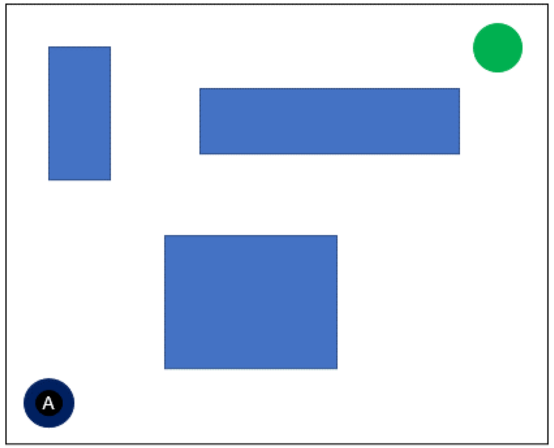
Reinforcement Learning

Aim of the work

Learning based path planning algorithm



Direct RL

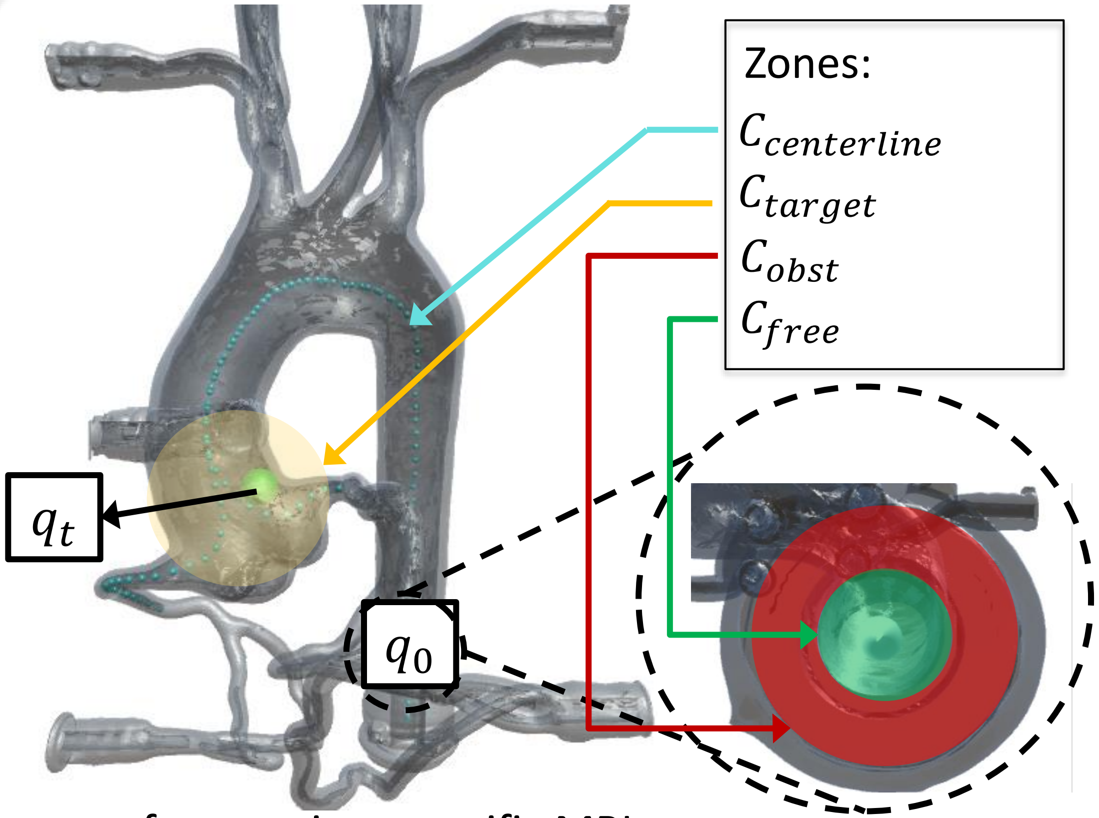


Inverse RL
Learning From Demonstrations

PCI navigation on 3D deformable environment
Replanning

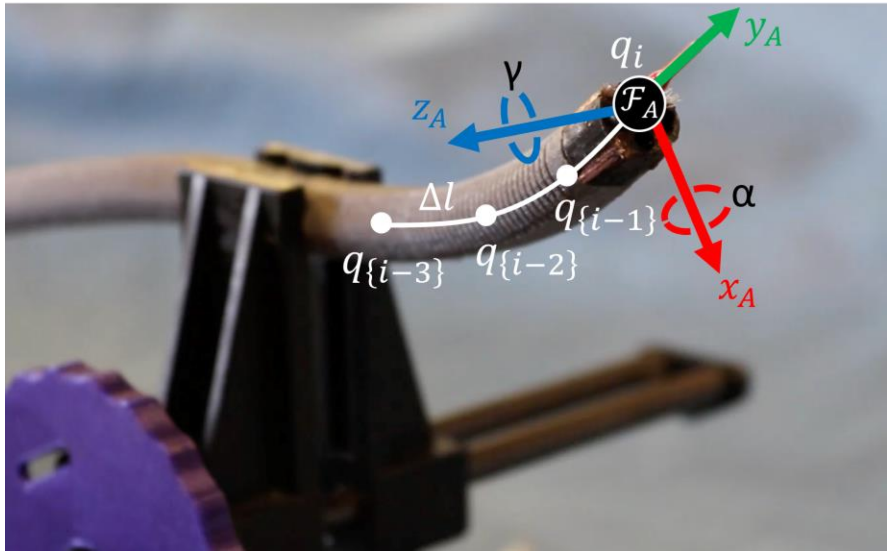
Modeling

DYNAMIC ENVIRONMENT



- from patient specific MRI
- position based dynamics

CATHETER

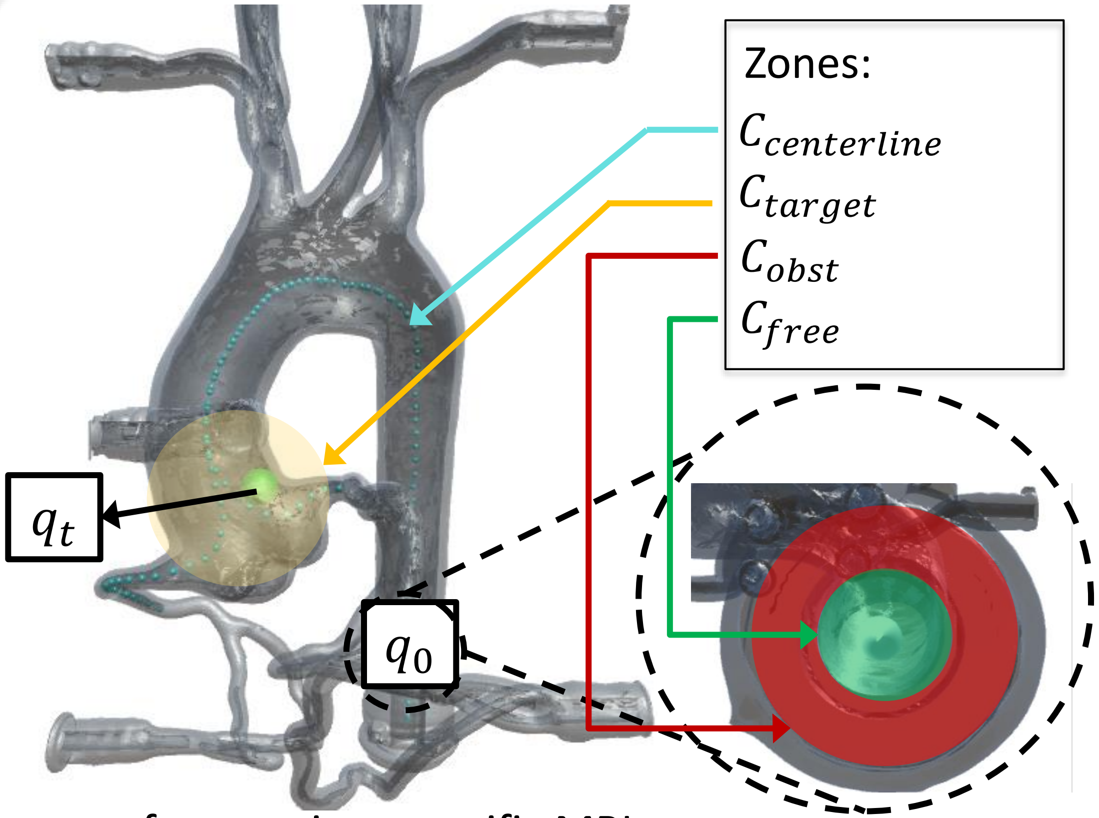


- Follow the leader model
- DOF: insertion, x-z rotation

[3] Ansari et al. (2022)

Modeling

DYNAMIC ENVIRONMENT



- from patient specific MRI
- position based dynamics

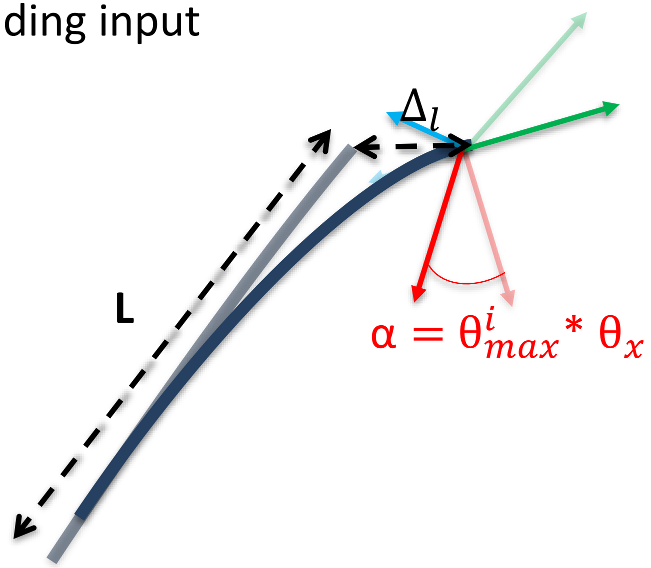
CATHETER

BENDING CONSTRAINT

$$\theta^i_{max} = \frac{\theta_{max} * \Delta_l}{L}$$

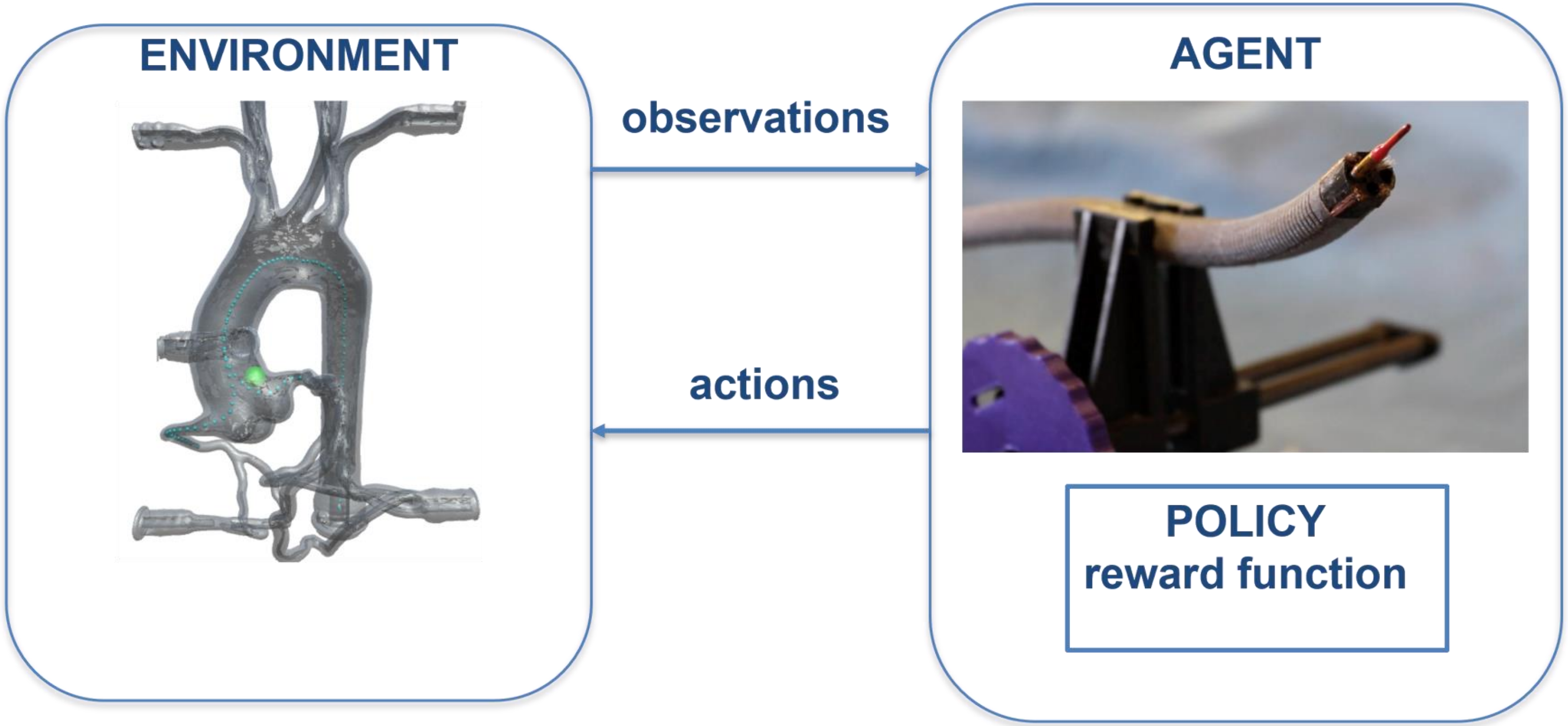
θ_{max}: maximum bending angle

θ_x : x-bending input



[3] Ansari et al. (2022)

Reinforcement Learning



Reinforcement Learning

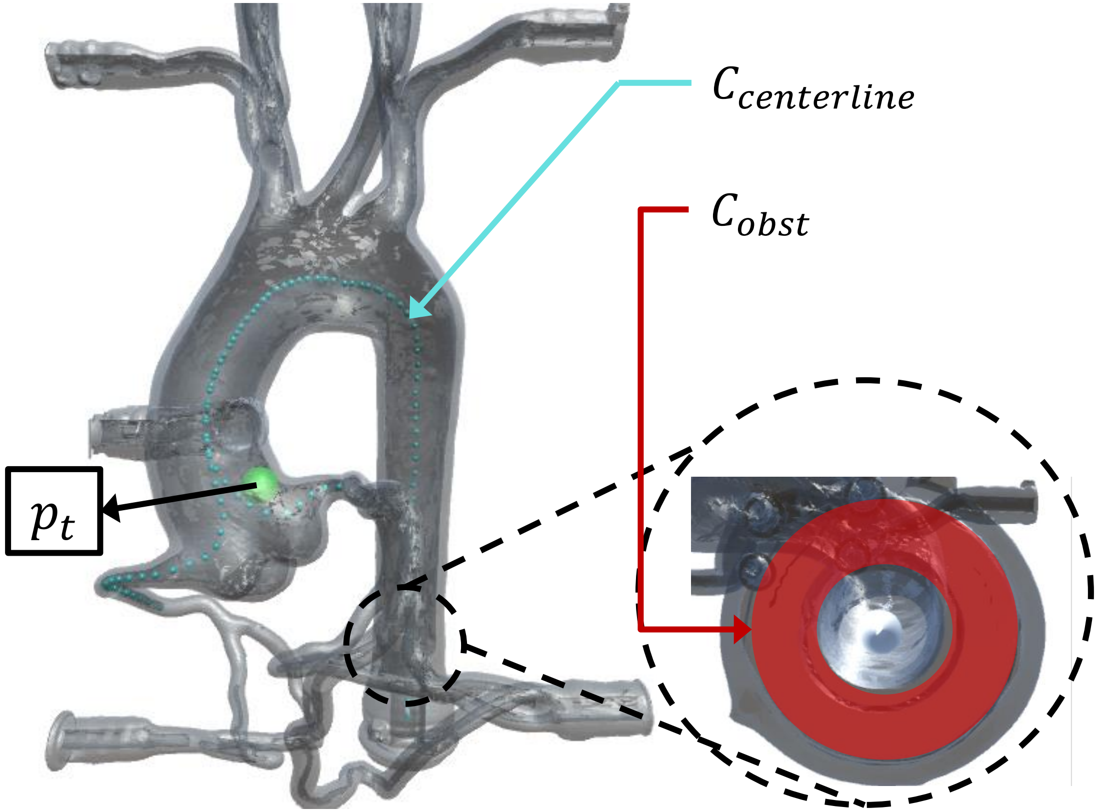
REWARD FUNCTION

$$r_t = r_{end} + r_{in}$$

$$r_{end} = \begin{cases} r_{obst} & if\ q_i \in C_{obst} \\ r_{target} & if\ \|p_i - p_t\| < \varepsilon \end{cases}$$

$$r_{in} = r_{step} + r_{centerline} + r_{bending}$$

r_{step}: time-step reward
$r_{bending}$: high angle reward



POLITECNICO MILANO 1863

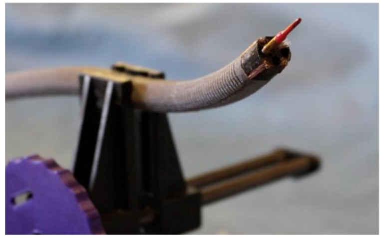
Reinforcement Learning

ENVIRONMENT

observations

actions

AGENT



POLICY
reward function

REWARD FUNCTION

$$r_t = r_{end} + r_{in}$$

$$r_{end} = \begin{cases} r_{obst} & \text{if } q_i \in C_{obst} \\ r_{target} & \text{if } ||p_i - p_t|| < \varepsilon \end{cases}$$

$$r_{in} = r_{step} + r_{centerline} + r_{bending}$$

ACTIONS

$$A = [\alpha, \gamma, \Delta_l]$$

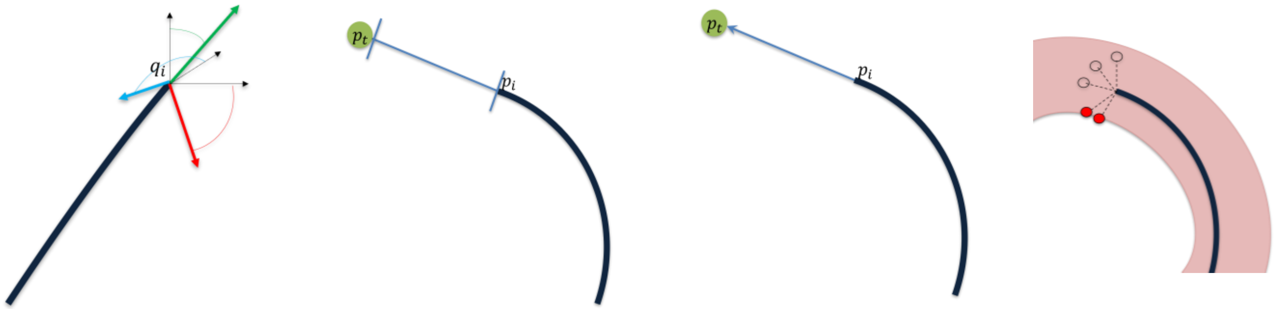
OBSERVATIONS

pose q_i

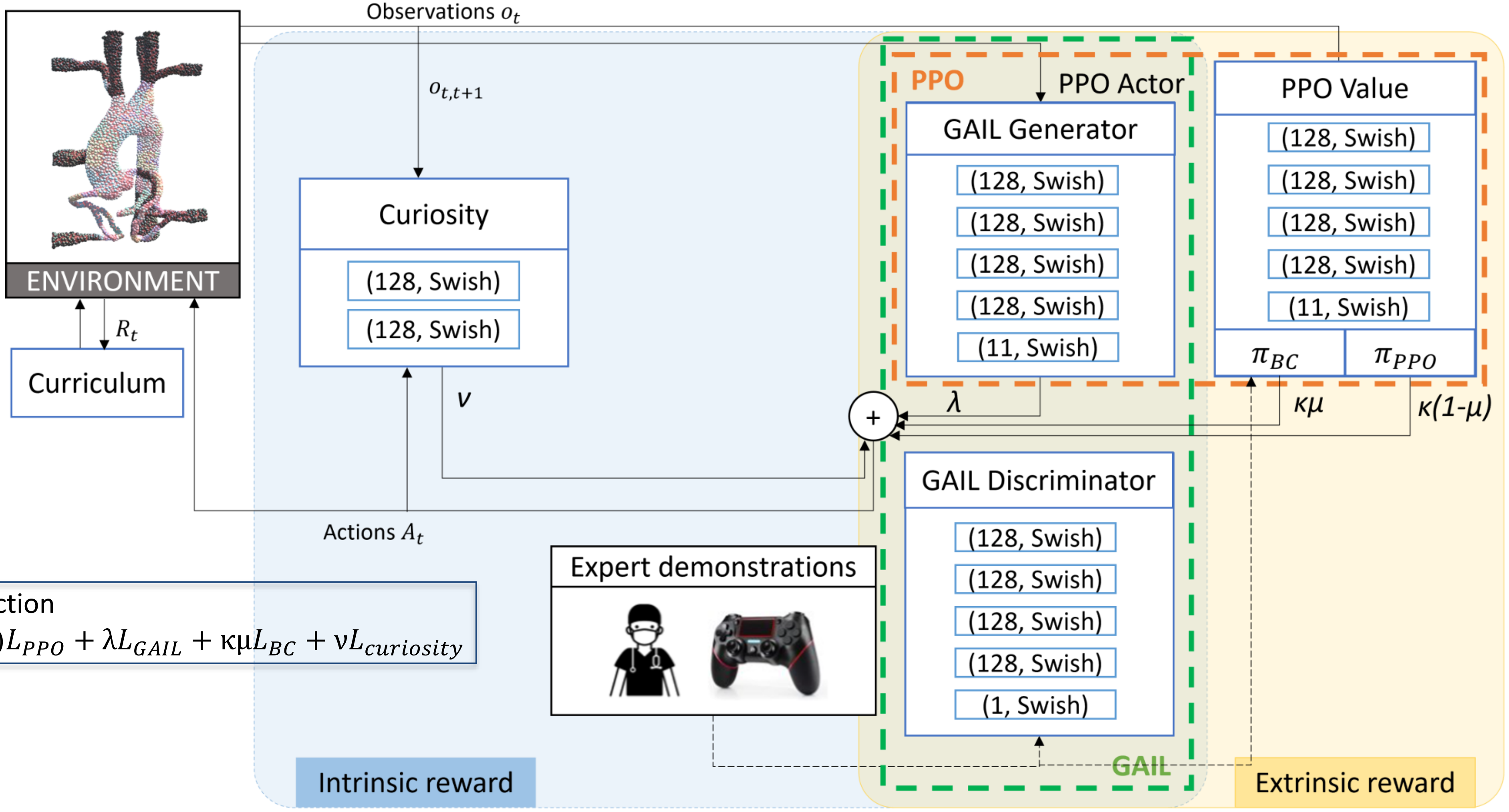
target distance $o_{dis} = \frac{||p_t - p_i||}{d_{max}}$

target direction $o_{dir} = p_t - p_i$

raycast o_{ray}



Network Architecture



Network loss function

$$L_{tot} = \kappa(1-\mu)L_{PPO} + \lambda L_{GAIL} + \kappa\mu L_{BC} + \nu L_{curiosity}$$

Performance metrics

Success rate

$\delta = n_s / n$

n_s: successful tests
n: total tests

Collision number

$$\varepsilon = \sum_{i=t_0}^{i=t_f} c(i)\, , c(i) = \begin{cases} 0, & if\ q_i \in C_{free} \\ 1, & otherwise \end{cases}$$

t_0, t_f: start time, stop time
$c(i)$: permitted collisions
q_i: agent pose at time i

Targeting error

$T_a = \min ||q_t - q_i||$

q_t: target pose

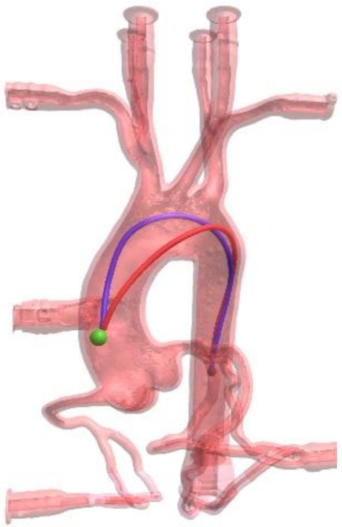
Duration

$T = t_f - t_0$

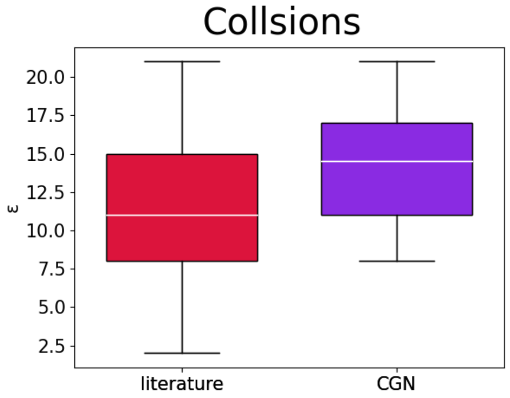
Experimental protocol

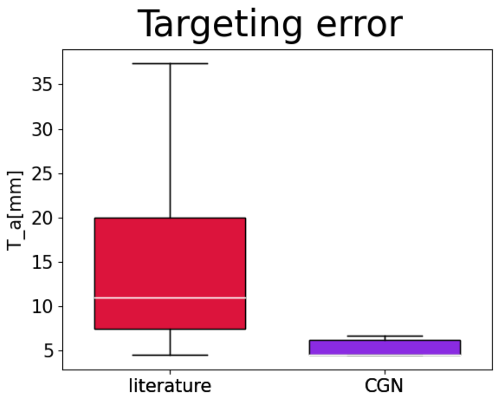
- 60 demonstrations from 1 expert (student)
- 5 million steps training
- 100 tests with the trained network
- 3 possible starting poses
- 5 possible target positions

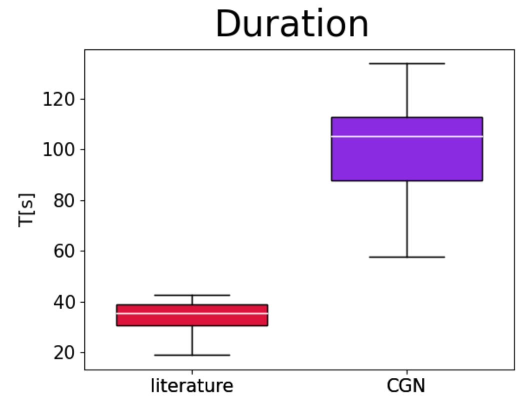
Comparison with literature



	δ
literature	14%
CGN	68%







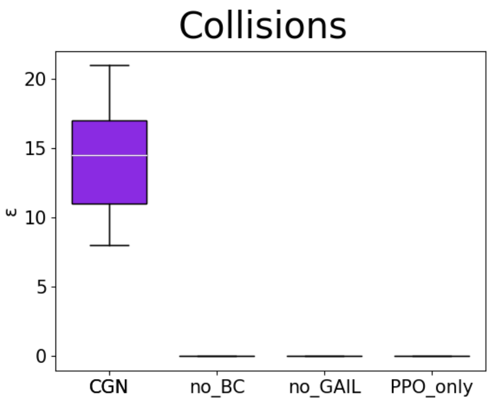
$p < 0.05$ using Kruskal-Wallis test

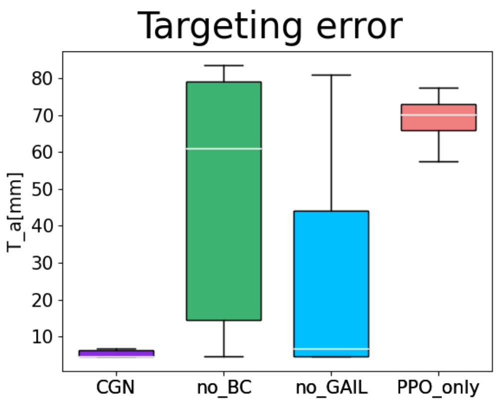
[4] Chi et al. (2020)

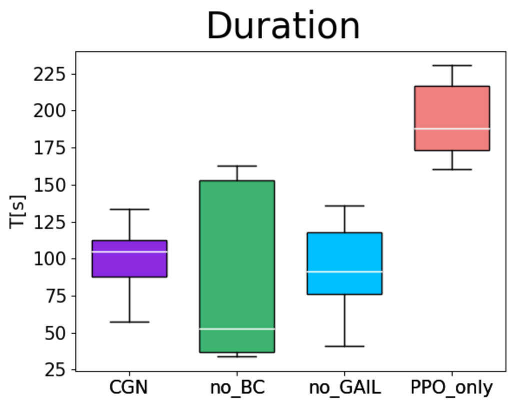
Ablation study

Ablation study

	δ
CGN	68%
no_BC	8%
no_GAIL	41%
PPO_only	0%







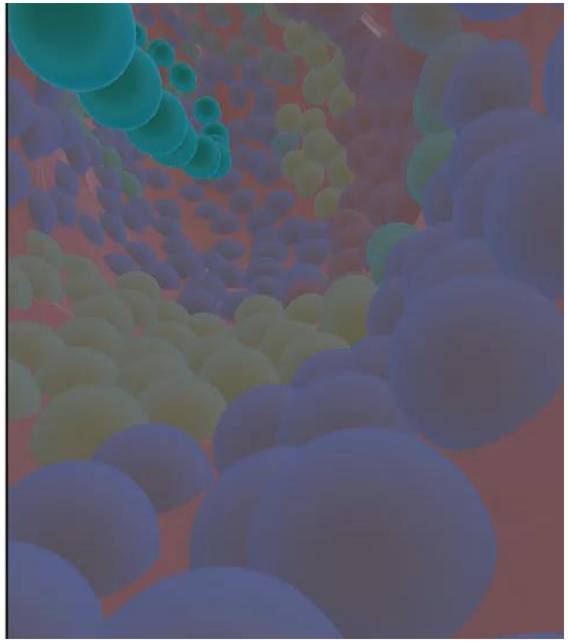
$p < 0.05$ using Kruskal-Wallis test

module	symbol	value
PPO	κ	0.2
BC	μ	0.7
GAIL	λ	0.8
curiosity	ν	0.02

Network loss function

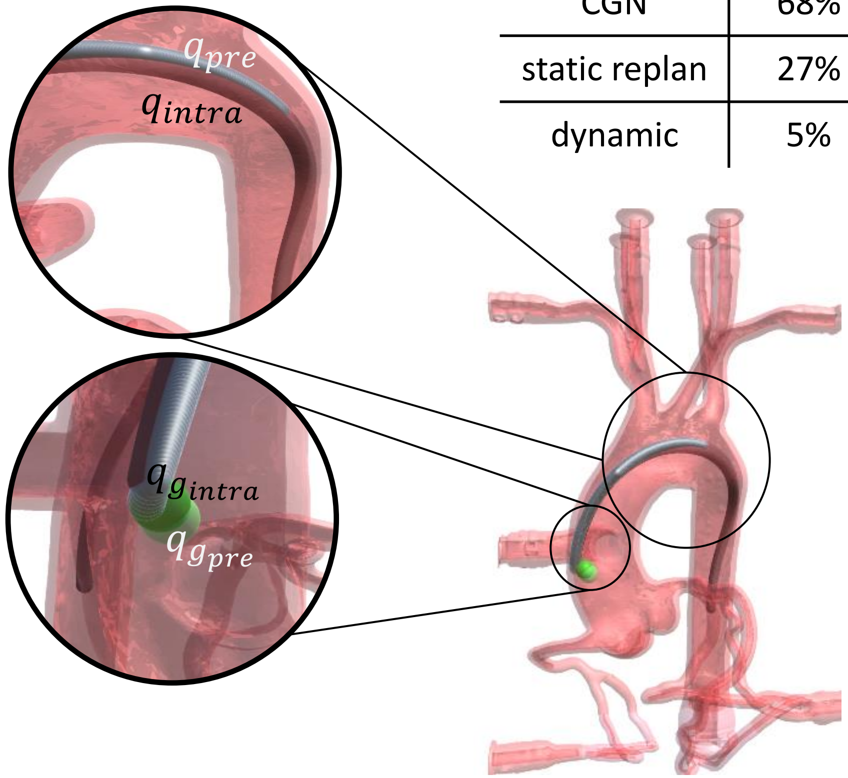
$$L_{tot} = \kappa(1-\mu)L_{PPO} + \lambda L_{GAIL} + \kappa\mu L_{BC} + \nu L_{curiosity}$$

Replanning

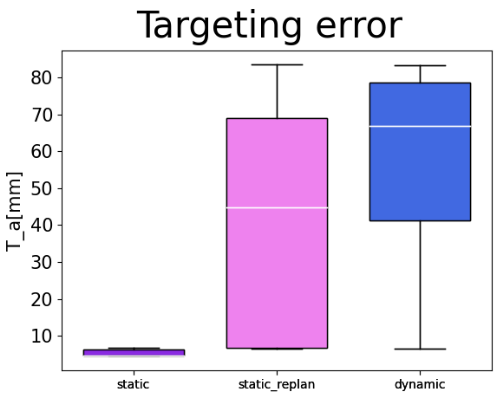


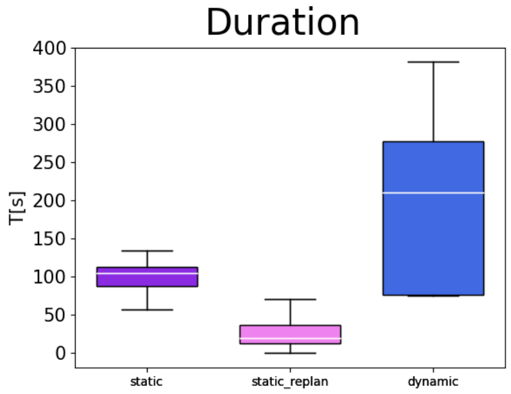
Graphical user interface:
external view and catheter view

Static and dynamic replanning



	δ
CGN	68%
static replan	27%
dynamic	5%





$p < 0.05$ using Kruskal-Wallis test

Conclusions and future work

Results

- Viable path and replanning
- Validation in a deformable environment
- Respect of physical constraints

Limitations

- Cumbersome model
- Limited precision for replanning

Future work

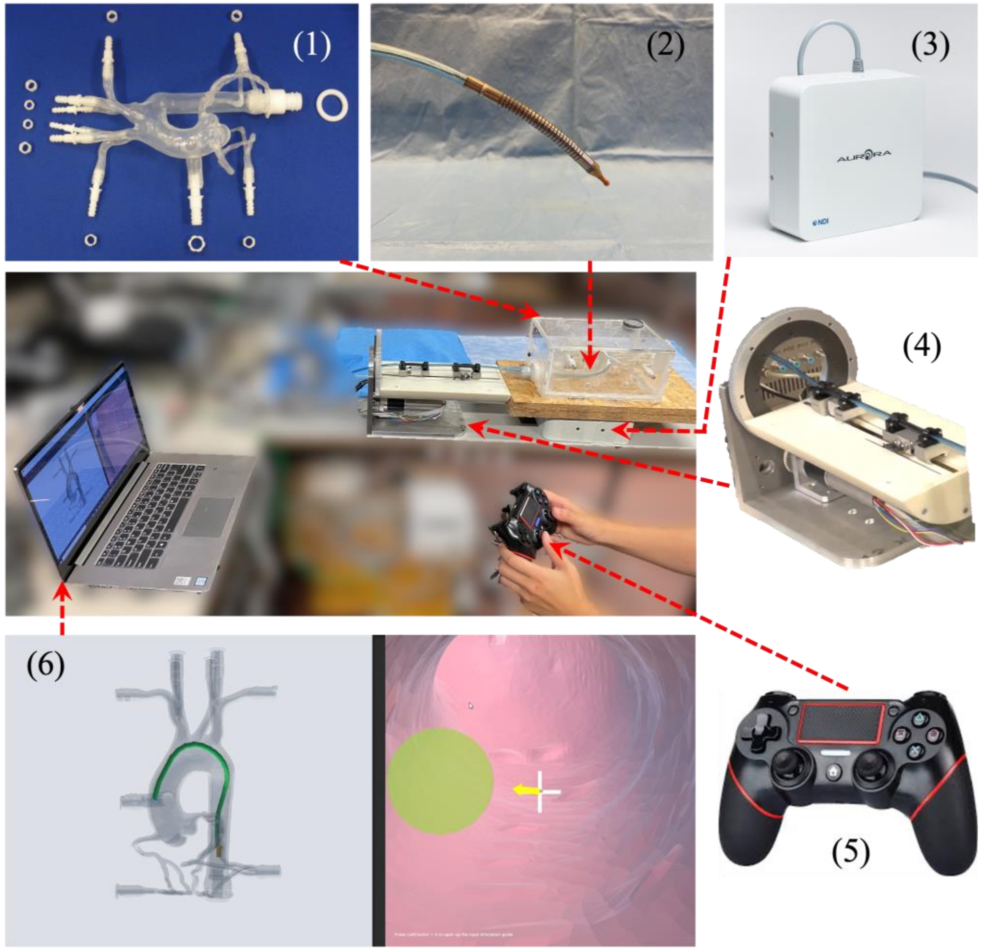
- Test on different aortic models
- Improve duration and precision of replanning

[5] Li Z., Lambranzi C. et al. Robust Path Planning via Learning from Demonstrations for Robotic Catheters in Deformable Environments, IEEE Robotics and Automation Letters

Thank you for your attention!
Questions?

Department of Electronics, Information and Bioengineering
Master of Science – Biomedical Engineering

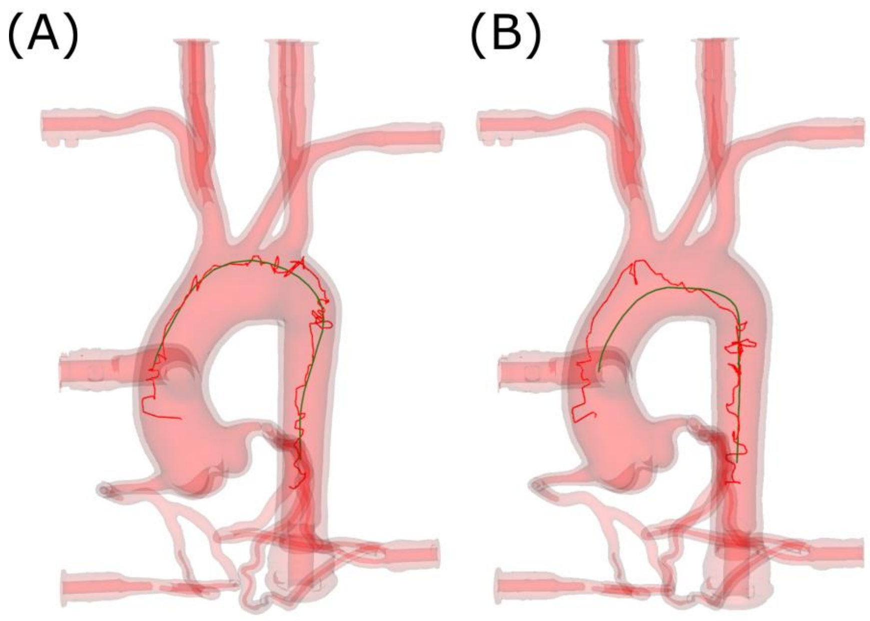
In vitro validation



SETUP

1) Silicone phantom
2) Catheter
3) EM sensor
4) Actuator
5) Controller
6) GUI
7) Provided path

RESULTS



(A) Tip position following the Curriculum GAIL path
(B) Tip position following the centerline