Inverse Reinforcement Learning Algorithm for Intra-Vascular and Intra-Cardiac Catheter's Navigation in Minimally Invasive Surgery

Andrea Fortuna, Mattia Magro, Angela Peloso, Alice Segato and Elena De Momi

September 27, 2022

Inverse Reinforcement Learning algorithm for intra-vascular and intra-cardiac catheter's navigation in Minimally Invasive Surgery

1st Andrea Fortuna
Politecnico di Milano
Milan, Italy

2nd Mattia Magro
Politecnico di Milano
Milan, Italy

3rd Angela Peloso
Politecnico di Milano
Milan, Italy

4th Alice Segato
Politecnico di Milano
Milan, Italy

5th Elena De Momi
Politecnico di Milano
Milan, Italy

***Abstract*—Structural Intervention Cardiology (SIC) is a mini-invasive intervention with a catheter based approach for cardiac surgery. Although SIC procedures are becoming increasingly popular, procedures are not ergonomic and technically demanding and, at the same time, high precision and accuracy in reaching target locations inside the human body are necessary for the success these procedures. Thus, there is therefore a need to develop a robust path planner framework to improve the accuracy in target reaching while minimizing interaction with anatomical structures. In this work a pre-operative path-planning method able to guide the catheter from the peripheral access to the desired target position with the needed orientation is proposed. The method exploits an Inverse Reinforcement Learning algorithm based on a combination of Behavioral Cloning (BC) and Generative Adversarial Imitation Learning (GAIL). The method was in-silico tested performing 50 intra-vascular and 70 intra-cardiac paths where the ratio between attempts in which the catheter reaches the target and total number of attempts, computation time, the difference between desired pose and the reached one were considered as validation metrics. Results show that the proposed method computes optimal path enabling the catheter to reach the target with an average error in position below 2 mm in the intra-vascular phase and below 1 mm in position and 6° in orientation in the intra-cardiac phase.**

***Index Terms*—Structural Intervention Cardiology, Imitation Learning, Steerable Needle, Surgical Planning**

I. INTRODUCTION

Structural Intervention Cardiology (SIC) is a medical specialty that focuses on therapeutic interventions for the heart cavities and vessels (coronary arteries) that do not require surgery, but a mini-invasive intervention with a catheter based approach. It minimizes the recovery time, risk of infection and it minimizes patient's trauma, while providing comparable efficacy as open chest surgery [1]. However, these procedures are not ergonomics and technically demanding so that the final outcome relies on the surgeon's skills [2]. Moreover, precision in catheter insertion and accuracy in reaching the target, while avoiding damaging anatomical structures, directly contribute to surgical operation outcome. In the last years, several research groups focused their efforts on the development of needles able to autonomously steer inside tissues. The advantage of these needles consist in the performance of curvilinear trajectories, planned to maximize the distance from sensitive anatomical structures to be avoided and maximize the precision in reaching the target [3], [4].
While conventional needles' insertion can be planned and performed by the surgeon, the complex kinematic constraints of an autonomous needle requires the aid of automatic or semiautomatic solutions for the insertion and navigation. Moreover, accurate placement of the needle tip inside tissue is challenging, especially when the target moves and anatomical obstacles must be avoided [5]. There is therefore a need to develop a robust path planner framework to improve the robustness in obstacle avoidance and risk management in complex environments that influence the insertion's accuracy from the peripheral access point to the intra-cardiac target point.
The current study focuses on the SIC treatment of Mitral Valve (MV) repair which provides a mechanical grasping of the damaged anterior and posterior leaflets of the MV using a small clip device, restoring MV coaptation. Thus, the aim of this work is to develop a pre-operative path-planning method to generate intra-vascular (IV) and intra-cardiac (IC) path safely and effectively from the peripheral access to the MV, where the clip of the catheter has to be implanted. This will enable the catheter to autonomously navigate to the anatomical region of interest, overcoming the errors introduced by the operator that could affect the accuracy in reaching target.
Several approaches for path planning have been proposed in literature such as graph-based and sampling-based methods. The former are based on the discrete approximation of the path planning problem and have the main advantage of being relatively simple to implement. On the other side, their computation time considerable increase as the simulation environment increases or becomes complex. The latter, instead, significantly reduces the computation time by sequentially sampling different points in the workspace and gradually constructs a data structure representing path without collisions. However, it may be the case that although a solution to the path planning problem is found using sampling-based methods, it doesn't coincide with the best and smoothest path to get to the target. To overcome these limits, in recent years Reinforcement Learning (RL) methods have experienced exponential growth due to their capability of solving path planning problems where traditional algorithm often fail. In these methods, an agent performs action within an environment and learns how to perform a specific task by receiving rewards from the

environment. In particular, in our work, a pre-operative path-planning is implemented exploiting an Inverse Reinforcement Learning (IRL) algorithm in which the agent learn rules by observing examples.

II. Materials and Methods

A. Moving Agent - Flexible Catheter

Let us consider the tip of the catheter as the agent of our application that at each timestep t can translate along the $z-axis$ and rotate about the x and y axes. The agent configurations are described by poses, denoted as 4×4 transformation matrices:

$$\mathbf{T}_{agent} = \begin{pmatrix} \mathbf{R}_{agent} & \mathbf{p}_{agent} \\ 0 & 1 \end{pmatrix} \tag{1}$$

where $\mathbf{p}_{agent} \in \mathbb{R}^3$ is the tip position and $\mathbf{R}_{agent} \in SO(3)$ is the orientation relative to a world coordinate frame.

B. Virtual environment

For our purposes, the simulation environment was represented by the anatomical structures of interest such as right and left atria and ventricles, inferior and superior vena cava, femoral veins, pulmonary artery (Fig.2) and was obtained from the segmentation and 3D reconstruction of a Computer Tomography (CT) scan. For the construction of the simulation environment we also defined:

- The "free space" C_{free} as the set of possible agent configurations;
- The "obstacle space" C_{obst} as the space occupied by obstacles;
- The "target space" C_{target} which denotes where we want the catheter to move to.

Although the application developed focused on the placement of a clip in the mitral valve, the flexibility of the implemented method allowed to adapt the algorithm according to each specific application by changing the geometrical properties of the catheter and the anatomical region of interest.

C. Pre-operative path planning

Minimally Invasive surgery aims at reaching the target safely and accurately. To do so the catheter has to minimize the interaction with the vessels' walls and avoid collision with the heart's walls. Therefore, it's necessary to plane a feasible path from the access outside the patient to the target inside the patient. Path planning algorithms create a geometric path in the simulation environment starting from the initial pose to the target one, while avoiding obstacles [6].
The path planner takes in input the starting configuration (q_s) of the agent, consisting in its pose (3 positions and 3 rotations in the 3 axes expressed in Euler angles) and the target configuration (q_t) and the output is a pre-operative path i.e. a sequence of a agent's configurations from the starting one to the target one "in press" [9].

D. Inverse Reinforcement Learning architecture

The proposed method exploited an Inverse Reinforcement Learning approach based on a combination of Behavioral Cloning (BC) [7] and General Adversarial Imitation Learning (GAIL) [8] achieving autonomous navigation inside the vessels and the heart to reach a goal location. The BC allows the agent to learn a specific task trying to mimic the behavior from demonstrations, whereas the latter deployed an adversarial approach using a discriminative next to a generative network "in press" [9]. At each time step (t) the discriminator network takes in input the expert (Q^{Manual}) and agent (Q^G) trajectories. Subsequently these two trajectories are compared, generating an intrinsic reward (r_{in}) that relies on a similarity score, updating the agent's policy (π). This loop carries on and it stops when the generator produces a path similar to the one from the expert's demonstrations.

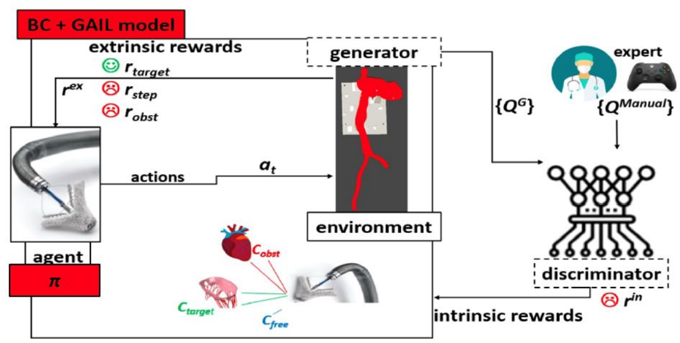


Fig. 1. Representation of the loop of the Inverse Reinforcement Learning algorithm based on the combination of Behavioral Cloning and Generative Adversarial Imitation Learning (GAIL); r_{target}, r_{step} and r_{obst} are part of environment's rewards defined in Sec. II-E.

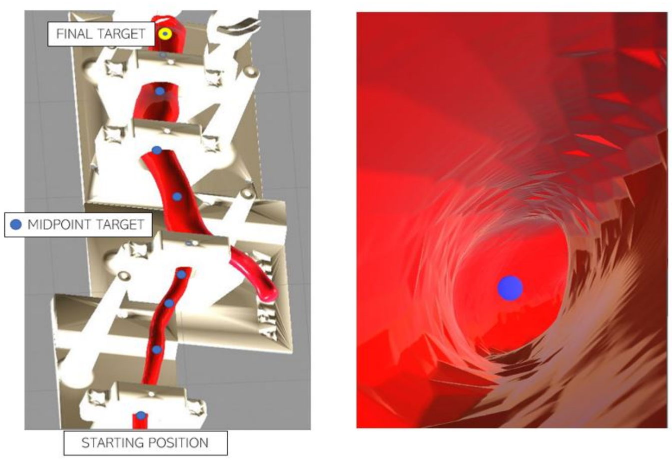
In addition, our catheter was placed inside an environment and it could take actions (a_t), moving towards the target. Depending on these actions, the environment could give positive or negative rewards, which were scalars, to the agent. This is named extrinsic reward (r_{ext}) and it was generated by the interaction between the agent, represented by the catheter, and the simulation environment giving a contribution in the development of the policy. Thus, the overall policy took in consideration both the information coming from the discriminator and from the environment.
Due to the complexity of the task and of the environment in which the agent has to move, the overall task was divided into two subsections: intra-vascular (IV) and intra-cardiac (IC) path.

E. Reward function

1) Intra-vascular path: The IV section covered the path from the peripheral access point of the catheter to the access of the heart's chamber, passing through the whole length of the femoral vein. Although the intrinsic function is the same for IV and IC, the extrinsic reward function was different for the two subsections. For the planning of the IV phase, the reward function has been shaped in order to make the agent learn to optimize the path, according to three main requirements:

- agent steps number minimization (t) using a negative reward, r_{step}, if the number of steps exceeded the predefined maximum number of steps allowed t_{max};
- minimization of interaction with vessels' walls using a negative reward, r_{obst}, if a collision was detected;
- Target Position Error (TPE) minimization, where TPE is the Euclidean distance between the needle's final position ($\mathbf{p}(\mathbf{q}_{agent})$) and the target position ($\mathbf{p}(\mathbf{q}_t)$), using a positive reward, r_{target}, if the target was reached.

Moreover, because of the complex geometry of IV path, a set of 7 intermediate target (Fig.2) were placed along the length of the vein in the training phase and to each of them a positive reward, r_{midtar}, was associated. This aspect was fundamental for the success of the training. Indeed, in our study, performing the IV training giving as input only the initial and final position did not allow the catheter to successfully find the path to the target.



(a) (b)

Fig. 2. (a) Intra-vascular environment represented by the femoral vein: between the starting and final point the representation of midpoint targets is shown; (b) midpoint target's depiction seen from the inside of the vessel.

The optimal parameters of the reward function for IV section, obtained with an empirical method, are reported in Table I.

TABLE I
REWARD FUNCTION PARAMETERS VALUES

path	r_{step}	r_{obst}	r_{target}	r_{midtar}	t_{max}
IV	-1	-1	+1	+0.08	10000

2) Intra-cardiac path: The IC section was represented by the path from the heart's access to the final target position. In addition to step minimization, obstacle avoidance and TPE, the Target Orientation Error (TOE), i.e. the difference in degrees between the z-direction of the target and the z-direction of the catheter, was also taken into consideration. Thus, the reward function was shaped according to the following requirements:

- agent steps number minimization (t) using a negative reward, r_{step}, if the number of steps exceeded the predefined maximum number of steps allowed t_{max};
- avoiding collision with heart's walls using a negative reward, r_{obst}, if a collision was detected;
- Target Position Error (TPE) minimization using a positive reward r_{target} was reached;
- Target Orientation Error (TOE) minimization using a negative reward r_{TOE}.

The orientation is crucial in clinical application for the correct placement of the catheter's clip. Indeed, the latter has to be placed in the correct position and with an orientation which, in the optimal condition, should be perpendicular to the plane of the MV (Fig.3). In order to achieve this aim, the target position was always chosen inside a plane parallel to the MV plane (Fig.3a) and in order to minimize the TOE, a negative reward, r_{TOE}, was added in the reward function of the IC section for every episode in which the TOE is greater than 10°. Finally, no intermediate targets were placed in the IC pathway.

The optimal parameters of the reward function for IC section, obtained with an empirical method, are reported in Table II.

TABLE II
REWARD FUNCTION PARAMETERS VALUES

path	r_{step}	r_{obst}	r_{target}	r_{TOE}	t_{max}
IC	-1	-1	+1	-1	5000

III. EXPERIMENTAL PROTOCOL

The IV model was trained on one single starting point, corresponding to the peripheral entry point, and tested on one single path. The episode (50 total episodes) was considered successful if the final position difference between the catheter pose and target one, i.e. TPE, was within 3.5mm. For the IV subsection, no restrictions were applied on the orientation of the catheter with respect to the target. The IC model, instead, was trained to reach any point placed in correspondence of the MV, starting from one single starting point. Thus, the test phase was conducted performing 70 episodes in which the target was randomly chosen inside the anatomical plane of the MV. In this case the episode is successful if the TPE was within 2 mm and the TOE was within 15°.

IV. VALIDATION

The proposed method was in-silico validated and for each of the pre-operative paths, in both intra-vascular and intra-cardiac environments, the quantitative analysis considered the following indexes:

- The success rate (SR[%]) i.e. the ratio between the insertions reaching the target and the total number of simulated insertions.

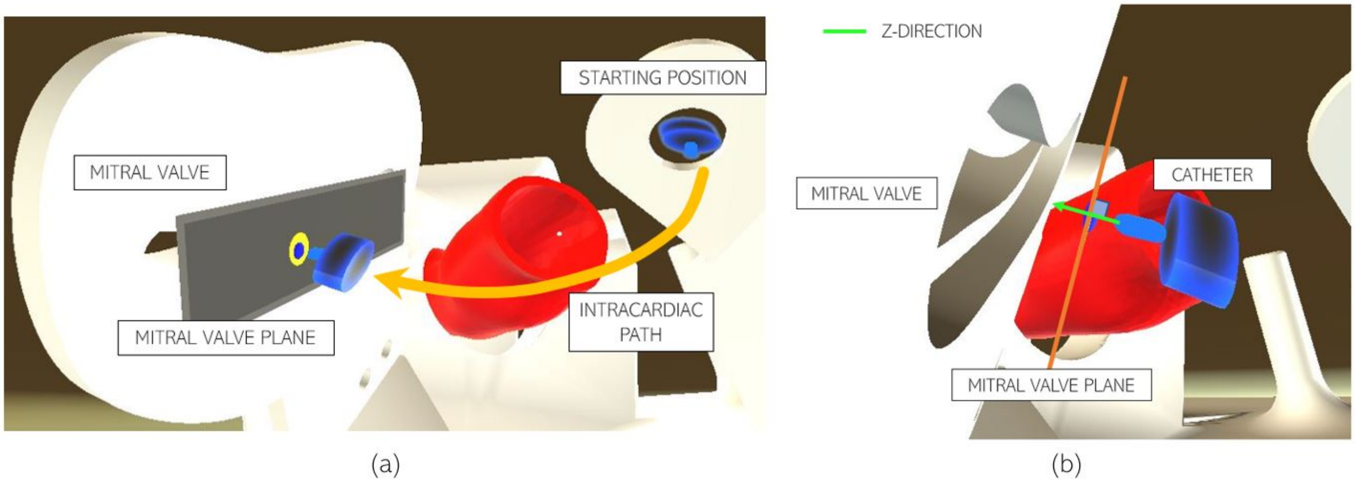


Fig. 3. (a) Representation of the Intra-cardiac path from starting to final position: the target point is randomly chosen inside the anatomical plane of the Mitral Valve (grey plane)); (b) in the optimal scenario, the angle between the tip of the catheter and the Mitral Valve plane has to be equal to 90°.

- Time (*time* [s]) required to perform the path, from the starting point to the target one.
- Target position error (TPE [mm]), i.e. the Euclidean difference between the needle's final position and the target position.
- (TOE [°]), i.e. the difference between the needle's final orientation and the target one, taken into consideration only during IC validation.

V. Preliminary Results

In the IV environment, the achieved success rate was 79 % with an average time needed to compute the path equal to 32.12±0.11 s and targeting position accuracy of 1.77±0.88 mm. In the IC environment, instead, the success rate was 88.6 % with an average time needed to compute the path equal to 10±1.2 s; the targeting orientation accuracy was 5.99±3.10° and the targeting position accuracy was 0.63±0.36 mm.

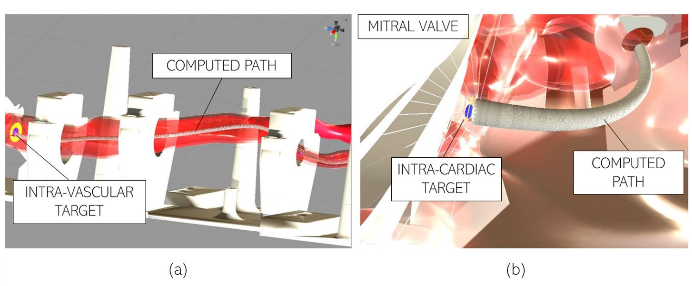


Fig. 4. (a) computed path for the intra-vascular section ; (b) computed path to the target placed upon Mitral Valve in the intra-vascular phase.

VI. Discussion and Conclusion

The presented work assessed the performance of a new IRL-based path planner for steerable needles able to minimize the interaction with vessel's walls during the intra-vascular navigation and avoiding collision with heart's anatomical obstacles during the intra-cardiac phase. The possibility of learning from a set of demonstrations, provided by an expert, allowed to consider all the requirements for the generation of the path that cannot always be included using graph- or sampling-based approaches. Thus, the proposed methodology can be exploited to calculate the optimal pre-operative path to the target. However, being an offline method, intra-operative information can not be used to update the optimal path based on real time data. Further perspectives may include the implementation of an online path re-planner able to re-compute in real time the path based on the information coming from the actual real position of the catheter, also considering the motion of anatomical structures.

References

[1] P. Lanzer, “Textbook of catheter-based cardiovascular interventions,” Cham: Springer International Publishing, 2018.

[2] A. K. Chhatriwalla, S. Vemulapalli, D. R. J. Holmes, D. Dai, Z. Li, G. Ailawadi, D. Glower, S. Kar, M. J. Mack, J. Rymer, A. S. Kosinski, and P. Soraja, “Institutional experience with transcatheter mitral valve repair and clinical outcomes: Insights from the tvt registry,” JACC Cardiovascular Interventions, vol. 12(14), pp. 1342–1352, 2019.

[3] J. Wang, T. Ohya,H. Liao, I. Sakuma, T. Wang, I. Tohnai, and T. Iwai, ”Intravascular catheter navigation using path planning and virtual visual feedback for oral cancer treatment”, The International Journal of Medical Robotics and Computer Assisted Surgery, 2011.

[4] A. Favaro, R. Secoli, F.R. Baena and E. De Momi, “Model-based robust pose estimation for a multi-segment, programmable bevel-tip steerable needle,” IEEE Robotics and Automation Letters, 2020.

[5] A. Segato, F. Calimeri, I. Testa, V.Corbetta, M. Riva and E. De Momi, “A hybrid inductive learning-based and deductive reasoning-based 3-D path planning method in complex environments,” Autonomous Robots, 2022.

[6] A. Gasparetto, P. Boscariol, A. Lanzutti, and R. Vidoni, “Path planning and trajectory planning algorithms: A general overview,” Motion and operation planning of robotic systems, pp. 3–27, 2015.

[7] T. K. Devi, A. Srivatsava, K. K. Mudgal, R. R. Jayanti, and T. Karthick, “Behaviour cloning for autonomous driving.” Webology, 2020.

[8] J. Ho and S. Ermon, “Generative adversarial imitation learning,” Advances in neural information processing systems, vol. 29, 2016.

[9] X. Zhang, M.C. Palumbo, F.Perico, M.Magro, A. Fortuna, T. Magni, E. Votta, A. Segato and E. De Momi, “Robotic Actuation and Control of A Catheter for Structural Intervention Cardiology”, IROS, 2016.