3D Neurosurgical Simulator for Training Robotic Steerable Catheter Agents Using Generative Adversarial Imitation Learning

1st Alice Segato
DEIB
Politecnico di Milano
Milan, Italy
alice.segato@polimi.it

2nd Irene Testa
DEIB
Politecnico di Milano
Milan, Italy
irene.testa@mail.polimi.it

3rd Elena De Momi
DEIB
Politecnico di Milano
Milan, Italy
elena.demomi@polimi.it

***Index Terms*—Robotic Steerable Catheter, Keyhole Neurosurgery, Generative Adversarial Imitation Learning, Neurosurgical Simulator, Inverse Reinforcement Learning**

I. Introduction

The complex anatomical structure of the brain and the vulnerability of its tissues make difficult to reach targets located in deep brain areas with rigid needles without causing damages to the adjacent tissues. Steerable catheter technology allow accessibility to the difficult anatomy [1]. Since a steerable catheter can only unfold its full potential when following the best trajectory within the brain, there is the need of pre-operatively planning the path of the probe minimizing length of the trajectory and maximizing distance to vessels, while respecting the kinematic constrains of the catheter. Among path planning methods for surgical steerable catheters, graph based and the sampling based methods, are unable in approaching human-level, generating a trajectory that does not take into account all the preferences of the expert surgeons. The approaches that may overcome these limitations are Learning Based methods as demonstrated by "in press" [1], making a system more robust since system parameters are adjusted automatically. In this work an imitation learning method is applied in which the agent learns to perform the desired trajectory thanks to a set of demonstrations given by an expert surgeon. The approach chosen to accomplish this task is the exploitation of Generative Adversarial Imitation Learning (GAIL) that is a technique that takes advantage of demonstrations executed by experts, and learns both the policy and reward function of the unknown environment [2]. This learning system, showed in Fig. 1, is applied to a 3D brain surgery simulator that has been developed with Unity and integrated with ML-Agents Toolkit [3].

This project has received founding from the European Union Horizon 2020 research and innovation program under grant agreement No 688279.

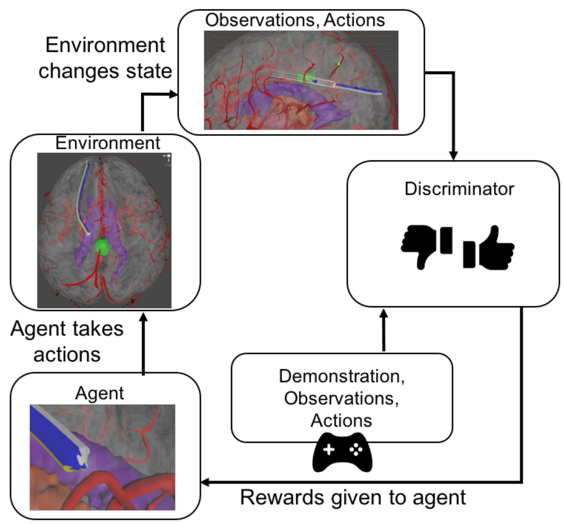

Fig. 1. Neurosurgical learning system using GAIL for Robotic steerable catheter.

II. Method

A. Environment Creation

In this work, preliminary tests were conducted on one case-study. High-resolution MR images including a 3D T1-weighted sagittal Fast-Field Echo and a 3D high-resolution time-of-flight MR angiography (TOF-MRA) were acquired. The two datasets: 1) ToF for vessels, 2) T1 for brain cortex were registered, allowing to obtain 3D models of the brain. The 3D brain surgery simulator used has been developed using Unity, the user can select the target (e.g. tumor) and the entry point of the catheter. At this point, the user can provide demonstrations, trying to reach the previously selected target with the use of the controller avoiding the obstacles, having the possibility to switch to an automatic path planner based on "Raycasts Algorithm" that explore the environment around the tip of the probe guiding it towards the target.

B. The Agent

This work is carried on in the context of EU Horizon EDEN2020 project which aims at advancing the current state of the art in neurosurgical technology, by a new biomimetic flexible probe (PBN) [4]. The agent in this work is the probe, it successfully learns to follow the best trajectory thanks to the observations that collects from the environment, and the rewards assigned to estimate the value of the agent's current state in accomplishing its tasks.

C. Observations

The observations of the environment are used by the policy to take a decision that is then forwarded to the agent. The observations are collected by "raycasts" and "spherecast" sensors attached to the agent. Through detectable "tags" it is possible to specify the types of objects that the agent should be able to distinguish between (i.e. the tumor and the anatomical obstacle). Hence, the set of rays and spheres that are cast allows the agent to explore the physics world in its moving direction with a wide visual field, and the anatomical structures that are hit determine the observation vector.

D. Reward Function

The main rewards in this model, used to obtain the optimal decision-making policy, encourage the reaching of the target without touching any anatomical obstacles along the followed path. As synthesized in the Algorithm 1, this is done by setting a positive reward if this happens and a negative one in the opposite case. Moreover, two negative rewards are added, one at each step of the catheter in order to obtain a faster execution of the task, and the other whenever the minimum distance to the obstacles is lower than a predefined distance.

Algorithm 1 Reward function - pseudocode for each episode

for step in maxstep **do**
  $moveAgent()$
  $mindist = GetNearestObstacleDist()$
  $addReward(-1/maxstep)$
  if $targetReached()$ is true **then**
    $addReward(1)$
    $newEpisode()$
  else if $collisionDetected()$ is true **then**
    $addReward(-1)$
    $newEpisode()$
  end if
  if $mindist < 1$ **then**
    $addReward(-0.1)$
  end if
end for

III. Results

A. Hardware Specification

We performed our experiments on a Linux machine equipped with a 6-core i7 CPU, 16GB of RAM and 1 NVIDIA Titan XP GPU with 12GB of VRAM.

TABLE I
Expert Demonstration vs Generated model trajectories

	Time[s]	Len[mm]	MinD[mm]	AvgD[mm]	MaxC[°]
Demo	7.251 ±3.41	64.75 ± 24.1	0.02 ± 0.06	9.73 ± 4.41	16
Model	2.038 ±0.73	55.65 ± 23.1	0.02 ± 0.03	11.78 ± 4.57	14

B. Experimental Protocol

The same experimental protocol was followed for both the expert demonstration and the GAIL generated model. At every episode a new starting point was randomly chosen on the right hemisphere and the target (the tumor) was moved in one out of 4 pre-selected position in order to encourage its ability to learn the correct policy, independently from the specific environment, starting point or target point.

C. Evaluation

Results, for both the expert demonstration and the generated model, are reported in Table I. The obtained trajectories were evaluated considering computational time length(Len), the minimum distance (MinD) and average distance (AvgD) from obstacles and the maximum curvature (MaxC).

IV. Discussion and conclusion

The machine learning methodology used in this work is imitation learning with demonstration. The agent was trained with a combination of GAIL and Behavioral Cloning (BC). The first one uses an adversarial approach to reward the agent for behaving similar to a set of demonstrations. While the second one trains the agent's neural network to emulate the actions shown in a set of demonstrations. In particular, the catheter represents the agent that is an entity that can observe its environment and handles generating its observations, performing the actions it receives and assigning a reward that can be positive or negative in order to obtain the best path inside the brain that reaches the target avoiding the obstacles. Results show a comparison between trajectories generated by the manual simulations and the generated model, highlighting the higher average distance from obstacles and the lower computational time, length and maximum curvature.

References

[1] A. Segato, L. Sestini, A. Castellano and E. De Momi "GA3C reinforcement learning for surgical steerable catheter path planning," 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2020.

[2] M. Chen, Y. Wang, T. Liu, Z. Yang, X. Li, Z. Wang, and T. Zhao "On computation and generalization of generative adversarial imitation learning," (2020).

[3] A. Juliani, V. P. Berges, E. Vckay, Y. Gao, H. Henry, M. Mattar, and D. Lange "Unity: a general platform for intelligent agents"(2018).

[4] C. Burrows, R. Secoli, and F. R. y Baena, "Experimental characterisation of a biologically inspired 3d steering needle," in 2013 13th International Conference on Control, Automation and Systems (ICCAS 2013). IEEE, 2013, pp. 1252–1257.