Jacqueline Ritter, Lennart Karstensen*, Jens Langejürgen, Johannes Hatzl, Franziska Mathis-Ullrich, and Christian Uhl

Quality-dependent Deep Learning for Safe Autonomous Guidewire Navigation

https://doi.org/10.1515/cdbme-2022-0006

Abstract: Cardiovascular diseases are the main cause of death worldwide. State-of-the-art treatment often includes the process of navigating endovascular instruments through the vasculature. Automation of the procedure receives much attention lately and may increase treatment quality and unburden clinicians. However, in order to ensure the patient's safety the endovascular device needs to be steered carefully through the body. In this work, we present a collection of medical criteria that are considered by physicians during an intervention and that can be evaluated automatically enabling a highly objective assessment. Additionally, we trained an autonomous controller with deep reinforcement learning to gently navigate within a 2D simulation of an aortic arch. Among others, the controller reduced the maximum and mean contact force applied to the vessel walls by 43% and 29%, respectively.

Keywords: deep reinforcement learning, safety, guidewire navigation, autonomous, machine learning

1 Introduction

One third of all deaths worldwide are caused by cardiovascular diseases. These include in particular the common maladies of heart attack and stroke, which are often treated with endovascular therapy. During an endovascular intervention, a catheter and a guidewire are inserted into an access vessel and steered through the vasculature until the target is reached. In order to navigate the endovascular device it is rotated and translated at the proximal end. During the procedure the surgeons are visually guided by 2D fluoroscopy images that show the position of the guidewire within the human vasculature.

***Corresponding author: Lennart Karstensen,** Fraunhofer IPA, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany, e-mail: lennart.karstensen@ipa.fraunhofer.de
Jacqueline Ritter, Jens Langejürgen, Fraunhofer IPA, Mannheim, Germany
Johannes Hatzl, Christian Uhl, Department of Vascular and Endovascular Surgery, University Hosptial Heidelberg, Heidelberg, Germany
***Corresponding author: Lennart Karstensen, Franziska Mathis-Ullrich,** Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Karlsruhe, Germany

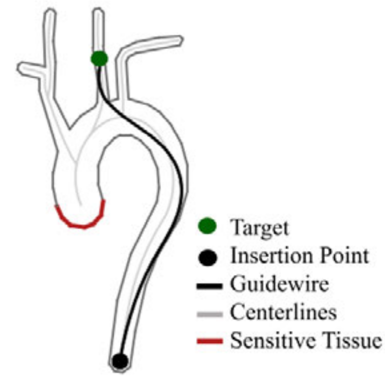


Fig. 1: Aortic arch model with centerlines showing the target, insertion point, the guidewire, and the sensitive tissue that leads to the heart valve.

Automation of this intervention has attracted much attention recently as it might improve patient safety and operation efficiency, reduce complications for the patients, and decrease fatigue and radiation of the clinicians. [1–3]

In order to assess the quality of a minimally invasive endovascular procedure experienced surgeons use structured grading scales such as the global rating scale (GRS) to ensure objective evaluation. The GRS rates the four main aspects of a procedure on a scale from one to five, i.e. the flow of operation, instrument handling, time and motion, and respect for tissue [4]. If the procedure is executed on a simulator, automatically measured metrics such as procedure time or the contrast volume can be used to evaluate the quality [5]. Rafii-Tari *et al.* [6] developed a framework to measure the catheter-tissue contact forces as well as operator motion patterns. Additionally, they indicate that a low standard deviation of the translation speed suggests smooth and controlled navigation behavior. In comparison to novices, experienced surgeons achieve a reduced number of translational guidewire movements, smoother motion in general, a shorter total path length of the device tip, and apply less torque and force to the device, which suggests that those criteria are an indicator for higher quality [7], [8].

In research regarding autonomous control of endovascular guidewires navigation quality is considered to a limited extend. Zhao *et al.* [1] train a CNN based controller with image and force measurements as input. The force data is used to detect a collision between the device tip and obstacles such as plague on the vessel wall. As a result, the controller then executes an avoiding action, i.e. pulling back the guidewire and rotating it by an angle. The system is trained on three different rigid models and tested on a separate one. The evaluation metrics comprise the target reached, the operating force, and the navigation time. Chi *et al.* [2] navigate a guidewire through a flexible aortic arch model by training a neural network with imitation and reinforcement learning. Electromagnetic tracking data of the device tip serve as input. The performance is evaluated on two different models and compared to the navigation of an experienced surgeon. The evaluation criteria comprise the path length, the mean and maximal force applied to the vessel walls, and whether the target was reached. Karstensen *et al.* [3] propose a controller that is trained in a simulation environment and evaluated in an ex-vivo specimen. The coordinates of the guidewire tip are used as input to train the neural network with deep reinforcement learning. Navigation quality is not considered, instead the controller is evaluated on the number of randomly distributed targets that can be reached within a fixed time span.

The aforementioned works mostly focus on whether the desired target point was reached and do not optimize for safe navigation. However, in order to keep the risk of damage as low as possible additional criteria need to be taken into consideration. In this work, we propose a collection of criteria that ensures the patient's safety during endovascular guidewire navigation. They can be measured automatically in a simulation environment resulting in a highly objective assessment of the quality. Additionally, we train an autonomous reinforcement learning controller that is able to reach arbitrary target points within a fixed aortic arch geometry and adheres to our criteria.

2 Methods

2.1 Gentle Navigation

The quality of the navigation process is evaluated with respect to the following criteria:

- Mean contact force applied to surrounding tissue.
- Maximum contact force applied to surrounding tissue.
- Total path length covered by device tip.
- Standard deviation of translation velocities.
- Number and distance of withdrawals.
- Total navigation time.
- Forward motion of device tip.
- Distance between tip and vessel centerlines.

Forward motion is defined as

$$forward\ motion = \frac{\Delta pos_{tip}}{v_{trans} \cdot \Delta_t} \tag{1}$$

where Δ_t denotes the duration of one control step, v_{trans} represents the translation velocity, and Δpos_{tip} defines the distance between the position of the device tip in the current and the previous control step. If the device tip moves freely it performs the same translation movement as the manipulation at its base and the value is close to one. Larger values are an indication that forces and torques are built up in the instrument, since the movement at the base is converted into deformation of the instrument instead of a movement of the tip. If the force and torque are released the instrument snaps back into its rest shape which increases the risk of perforation.

In order to avoid inaccuracies when computing the distance between the tip and the centerline we assume that the centerline points are densely sampled. For each control step the distance between the device tip and the closest point on the centerline is computed.

Adhering to the defined criteria ensures a high quality navigation process that minimizes the risk for complications for the patient. Furthermore, all criteria can be evaluated automatically, which makes the evaluation of the procedure highly objective and efficient.

2.2 Aortic Arch Environment

Training and evaluation of the controller are carried out in a 2D simulation environment. The shape of the aortic arch environment and the guidewire are shown in Fig. 1. The guidewire is modeled as a multibody system, in which the elements are connected by damped linear and damped rotational springs. The number of elements varies according to the total length of the guidewire. It is navigated through the aortic arch by translation and rotation at its base. The simulation itself is based on pymunk, a 2D rigid body physics library [9].

2.3 Autonomous Controller

The design of the neural network controller is based on the work of Karstensen *et al.* [3]. However, we enhance the feedforward architecture using the deep reinforcement learning algorithm soft actor critic (SAC), which is less sensitive to the correct choice of hyperparameters [10]. The output of the actor network consists of the mean and the standard deviation of a normal distribution for every dimension in the action space. The actions, i.e. translation and rotation velocities, are sampled from the corresponding distribution. The observations that serve as input to the neural network controller are defined as the current and last position of the guidewire tip, the action between them, as well as the target position. The targets are randomly distributed within the aortic arch excluding the aorta.

During training the navigation task of steering the guidewire from a starting point to a target is executed for $2 \cdot 10^6$ steps. One navigation task is called episode. An episode is completed either if the target point is reached within a threshold radius of 5 mm or 200 simulation steps have been performed, which corresponds to 27 seconds in a similar real world scenario.

Two different controllers were trained. The baseline controller is trained with an extended version of the sparse reward proposed in [3], where additionally the pathlength difference between two steps is taken into account. Pathlength is the distance between device tip and target along the centerlines. The reward extension provides the controller with further information about the environment that speeds up the training process when using SAC. The reward that is used to train the baseline controller is denoted as

$$R_{base} = -0.005 - 0.001 \cdot \Delta pathlength + \begin{cases} 1.0 & \text{target reached} \\ 0.0 & \text{else} \end{cases}$$

In order to train a controller that adheres to the criteria stated in Section 2.1 the reward R_{force} and R_{valve} are added to R_{base}. Note that since the criteria depend on each other it is sufficient to consider a subset for the reward function.

$$R_{force} = -4.93 \cdot 10^{-7} \cdot force$$

$$R_{valve} = \begin{cases} -0.1 & \text{device touches heart valve} \\ 0.0 & \text{else} \end{cases}$$

Here, the reward R_{force} punishes contact forces applied to the vessel wall. Note that its value is zero if no contact force is executed. R_{valve} penalizes any contact of the device with the part of the aortic arch that leads to the heart valve, which consists of particularly sensitive tissue. The area that models the sensitive tissue leading to the heart valve is highlighted in red in Fig. 1. The overall reward per step is denoted as $R_{quality}$ and combines the rewards described above.

$$R_{quality} = R_{base} + R_{force} + R_{valve}$$

2.4 Evaluation

Every $5 \cdot 10^4$ steps the success rate of the controller is evaluated for 1000 consecutive episodes. During evaluation, the mean of each distribution is used as an action, thus rendering it deterministic. Especially in medical applications the decision of an autonomous system needs to be traceable in order to prevent the controller from executing unforeseen behavior. The success rate is defined as the percentage of evaluation episodes where the controller is able to reach the target.

For each reward set-up the state of the trained controller that reaches the highest success rate is additionally evaluated on the criteria from Section 2.1 for 1000 episodes. Per episode the total path length is divided by the length of the optimal path along the centerlines in order to allow for comparison. The metric is denoted by the path length ratio. Its value should ideally be close to 1.0. The distance of withdrawals, the forward motion of the device tip, and the distance between the tip and the centerlines are averaged over all obtained values. For all other criteria the final value of the episode is used. Note that the simulation focuses on optimizing realistic behavior rather than realistic forces.

3 Results

The learning curve for the two controllers is depicted in Fig. 2. After $1.25 \cdot 10^6$ steps the quality controller reaches 85% of the targets. It continues to slowly improve until it reaches a maximum of 95.6% after $1.90 \cdot 10^6$ steps. The success of the baseline controller raises at a faster rate and reaches 89% after $0.35 \cdot 10^6$ steps and a maximum of 96.3% after $1.95 \cdot 10^6$ exploration steps.

The evaluation of the criteria from Section 2.1 is summarized in Tab. 1. The quality controller outperforms the baseline for most of the criteria. However, the baseline controller navigates closer to the centerlines, withdraws the guidewire less often and navigates quicker. Fig. 3 shows the trajectories of the two controllers in the aortic arch model aiming to reach the same target. The trajectory of both controllers have a similar length, while the baseline controller applies more contact force overall.

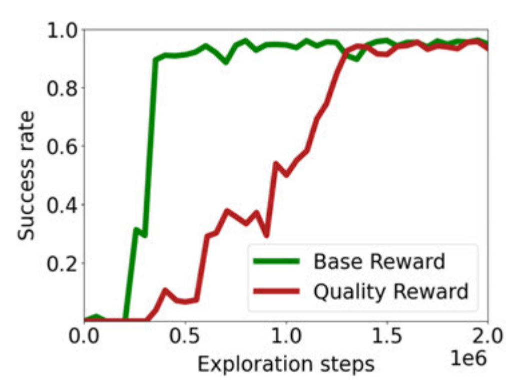
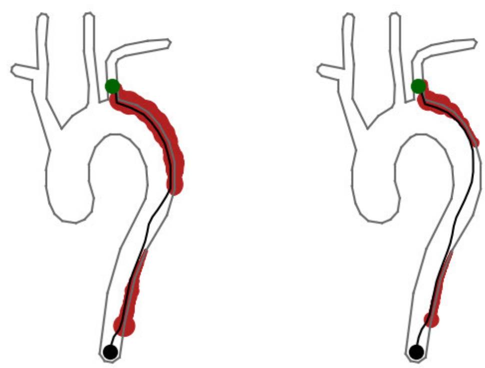


Fig. 2: Success rate of the baseline and quality controller during the training process.

Tab. 1: Evaluation of the controllers on the collection of criteria.

Metric	Baseline	Quality
Success rate	96.3%	95.6%
Mean contact force [mN]	6.81	4.83
Max. contact force [mN]	27.84	15.83
Path length ratio	1.09	0.98
Std translation [mm/s]	2.10	1.58
Number withdrawals	5.20	13.59
Withdrawal distance [mm]	16.47	8.50
Total navigation time [s]	6.76	7.07
Forward motion	1.20	1.01
Distance centerlines [mm]	4.49	4.76



(a) Baseline controller **(b)** Quality controller

Fig. 3: Navigation trajectories from the insertion point to the same target. Higher contact forces are indicated by a larger radius of the corresponding circle.

4 Discussion & Conclusion

The quality controller navigates the guidewire with a lower mean and maximum contact force, which decreases the risk for vessel damage for the patient. Fig. 3 shows the different behaviors when navigating the same target. The reduced contact force can be explained by the low standard deviation of the translation velocity resulting in a smoother and more controlled behavior. Additionally, when touching the wall and thus applying contact force the quality controller withdraws the device, rotates it and advances it rather than rotating the device while it pushes against the wall. Hence, the guidewire applies less force overall, but requires an increased amount of withdrawals and more time to reach the target. Consequently, the forward motion of the quality controller is closer to 1.0 corresponding to less build up of torque and force, and therefore less snapping.

Despite the good results for the simplified 2D simulation environment, the controller is yet to be transferred to a more realistic 3D environment and the real world, which will increase the complexity of the task and require hyperparameter adjustment. Using our approach the controller learns to navigate the unique aortic arch it is trained on. However, it needs to be able to generalize to unseen aortic arch geometries in order to make real world application possible. Future work should address this problem, e.g., by incorporating recurrent neural networks.

In conclusion, we derived criteria for the quality of guidewire navigation and successfully trained a deep-learning-based controller in a 2D simulation to improve these criteria compared to the state of the art.

Author Statement

This project is funded by the Ministry of Economics, Labor and Tourism Baden-Württemberg within the framework of the Forum Gesundheitsstandort Baden-Württemberg.

References

[1] Zhao Y, Guo S, Wang Y, Cui J, Ma Y, Zeng Y, et al. A CNN-based prototype method of unstructured surgical state perception and navigation for an endovascular surgery. robot. Med. Biol. Eng. Comput. 2019;57:1875-1887.

[2] Chi W, Liu J, Rafii-Tari H, Riga CV, Bicknell CD, Yang G Learning-based endovascular navigation through the use of non-rigid registration for collaborative robotic catheterization. Int. J. Comput. Assist. Radiol. Surg. 2018;13:855-864.

[3] Karstensen L, Ritter J, Hatzl J, Pätz T, Langejürgen J, Uhl C, et al. Learning-Based Autonomous Vascular Guidewire Navigation without Human Demonstration in the Venous System of a Porcine Liver. Int. J. Comput. Assist. Radiol. Surg. Forthcoming 2022.

[4] Reznick R, MacRae H Teaching surgical skills changes in the wind. N. Engl. J. Med. 2006;25:2664-2669.

[5] Mazomenos EB, Chang P, Rolls A, Hawkes DJ, Bicknell CD, Vander Poorten E, et al. A survey on the current status and future challenges towards objective skills assessment in endovascular surgery. J. Med. Robot. Res. 2013;0:1-19.

[6] Rafii-Tari H, Payne CJ, Liu J, Riga CV, Bicknell CD, Yang G Towards Automated Surgical Skill Evaluation of Endovascular Catheterization Tasks based on Force and Motion Signatures. IEEE Int. Conf. Robot. Autom. 2015;1789-1794.

[7] Rafii-Tari H, Payne CJ, Bicknell CD, Kwok K, Cheshire NJW, Riga CV, et al. Objective Assessment of Endovascular Navigation Skills with Force Sensing. Ann. Biomed. Eng. 2017;45:1315-1327.

[8] Rolls AE, Riga CV, Bicknell CD, Stoyanov DV, Shah CV, Van Herzeele I, et al. A Pilot Study of Video-motion Analysis in Endovascular Surgery: Development of Real-time Discriminatory Skill Metrics. Eur. J. Vasc. Endovasc. Surg. 2013;45:509-515.

[9] Blomqvist V Pymunk: An easy-to-use pythonic rigid body 2d physics library (version 6.2.1). 2007.

[10] Haarnoja T, Zhou A, Hartikainen K, Tucker G, Ha S, Tan J, et al. Soft Actor-Critic Algorithms and Applications. arXiv:1812.05905 2019.