

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/360758140>

Surgical GAN: Towards Real-time Path Planning for Passive Flexible Tools in Endovascular Surgeries

Article in *Neurocomputing* · May 2022

DOI: 10.1016/j.neucom.2022.05.044

CITATIONS

0

READS

61

8 authors, including:



Yan Zhao

Beijing Institute of Technology

34 PUBLICATIONS 420 CITATIONS

SEE PROFILE



Yuxin Wang

Beijing Institute of Technology

14 PUBLICATIONS 219 CITATIONS

SEE PROFILE



Jianhua Zhang

Hebei University of Technology

147 PUBLICATIONS 1,742 CITATIONS

SEE PROFILE



Xinke Liu

Beijing Tiantan Hospital

25 PUBLICATIONS 323 CITATIONS

SEE PROFILE

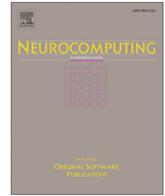
Some of the authors of this publication are also working on these related projects:



Amphibious spherical robots [View project](#)



Catheter Minimally Invasive Surgery Robot System, Force Feedback [View project](#)



Surgical GAN: Towards real-time path planning for passive flexible tools in endovascular surgeries



Yan Zhao^a, Yuxin Wang^b, Jianhua Zhang^{a,*}, Xinke Liu^d, Youxiang Li^d, Shuxiang Guo^{b,c}, Xu Yang^a, Shunming Hong^e

^aSchool of Mechanical Engineering, Hebei University of Technology, Tianjin 300132, China

^bSchool of Life Science, Beijing Institute of Technology, Beijing, China

^cFaculty of Engineering, Kagawa University, 2217-20 Hayashi-cho, Takamatsu, Kagawa, Japan

^dDepartment of Interventional Neuroradiology, Beijing Neurosurgical Institute and Beijing Tiantan Hospital, Capital Medical University, Beijing, China

^e3rd Medical Center of Chinese PLA General Hospital, Beijing, China

ARTICLE INFO

Article history:

Received 7 April 2021

Revised 22 February 2022

Accepted 14 May 2022

Available online 21 May 2022

Communicated by Zidong Wang

Keywords:

Surgical path planning

Endovascular surgery

Surgical GAN

Passive flexible tool

ABSTRACT

Automatic surgical path planning of the passive flexible tool encounters a prohibitive challenge, typically in endovascular surgery (ES). The key problem is that unstructured surgical environment and tools' unpredictable motion is hard to be explicitly modeled. We propose a generative adversarial networks (GAN)-based framework (defined as surgical GAN) towards automatic guidewire path planning in real time for ES. A novel GAN architecture is proposed by combining convolutional neural networks (CNN) and long short-term memory networks (LSTM), which extracts and fuses the spatial features in medical images and temporal features of historical tool path as the conditional information. It inputs the surgical state information and continuously outputs the local future path of the guidewire tip. A training dataset (3.5×10^5 samples) is collected under laboratory conditions with 10 surgeons. Effects of different CNN architectures and path planning length on network performance are investigated. User experiments, with the tasks delivering the guidewire tip inside a vascular model (endovascular evaluator) from the aortic arch into the left common carotid artery (LCCA), left subclavian artery (LSCA), or brachiocephalic trunk, are conducted with 10 novice surgeons in an operating room. The results shows significant improvement of a path planning accuracy with surgical GAN compared with baseline networks (from 46.2%–69.78%) and the non-rigid registration method (72.94%). Results of user experiments demonstrate an overall better task performance with the guidance of planned surgical path. Collectively, surgical GAN can achieve real-time path planning of the guidewire in simulated ES, and holds great potential for novice training and robotic ES autonomy.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Even the most experienced surgeon may still experience difficulties in finding the optimal surgical path for each surgery [1]. The main reasons are the variability of anatomical structure between patients and the complexity of the surgical environment. Computer-assisted automatic surgical path planning is beneficial for this problem and has potential advantages as follows, but it remains a challenge when faced with passive flexible tools, of which the motion is hard to be modeled and predicted in real time, in soft tissue surgery like the endovascular surgery (ES).

(1) Computer planned surgical path can be obtained by combining the experience of more than one expert surgeon. Their previous experience can be embedded in the difficult surgical conditions[2].

(2) During the training procedure, the novice surgeon can practice under the guidance of optimal planned surgical paths, which is conducive to the novice surgeon to learn the expert surgeons' experience[3].

(3) The autonomy of current master-follower surgical robots [4,5] can be enhanced by autonomously tracking the planned surgical path. It has been demonstrated that the semi-autonomous robot assisted surgery provides potential access to optimal surgical outcomes [2].

The unstructured surgical environment causes difficulties for surgical path planning. Human tissues are always of irregular

* Corresponding author at: School of Mechanical Engineering, Hebei University of Technology, No. 5340 Rd. Xiping Beichen District, Tianjin 300401, PR China.

E-mail address: jhzhang@hebut.edu.cn (J. Zhang).

shapes, the vessel for instance. In practice, it is hard to explicitly model the anatomical shape of the vessel in real time. In several previous researches of path planning for ES and ES robots [6,7], the extracted vascular center line was regarded as the surgical path for delivering the guidewire tip. But it is obviously unfitting since the guidewire advances always along the vascular wall rather than the center line. In addition, during guidewire insertion, the surgeon always intermittently injects a low dose of contrast agent under X-ray to observe the local surgical environment via 2D angiography images. Only the varying local vascular image can be seen in real time rather than the global image [8]. So, a real-time local path planning method is important for ES, which is capable of modeling the unstructured and changeable local scenes in 2D angiography images.

The path planning in ES would be more difficult, when a passive flexible surgical tool (like the guidewire commonly used in ES) is used, because it is soft and its motion is hard to be modeled and predicted in real time. Surgical path planning for rigid instruments including puncture needles [1] and even controllable flexible guidewires [9] are convenient, because of their available kinematic models. However, the motion of a passive flexible tool is hard to be modeled and calculated in real time due to its unpredictable deformation. Although the finite element model method has potential to model the guidewire, it is time consuming and always used for off-line analysis [10].

Both the unstructured and changeable property of local scene and unpredictable deformation of the passive flexible tools are essentially highly non-linear problems. An explicit mathematical model is hard to be established for surgical path planning in such surgeries due to the high nonlinearity. Recently, deep learning attracted wide attention on medical image processing [11–13] and showed outstanding capability for solving high-nonlinear problems. In our previous work [14], a trained convolutional neural network (CNN) is directly used for automatic navigation of basic guidewire insertion tasks in a Y-shaped vascular model. It indicates that deep learning provides a promising way to deal with the above challenge.

In this paper, we present a deep learning-based framework to address the challenge for real-time planning of guidewire path in ES. We propose a novel Generative Adversarial Networks (GAN) architecture, which is defined as surgical GAN. The proposed framework is capable of modelling the unstructured surgical environment and motion of passive flexible tools.

Our main contributions are as follows:

- (1) We propose a novel surgical GAN architecture, which has the capability of extracting and fusing spatial features and temporal features. It is enabled by constructing the generator and discriminator using combined CNN and LSTM.
- (2) We present the attempt to employ GAN for surgical path planning of a passive flexible tool in ES. It employs the features extracted from both preprocessed medical images and historical tool path as the conditional information, and outputs in real-time the local future path of guidewire tip.
- (3) Comparing with baseline networks and the non-rigid registration method, we demonstrate the advantage of the proposed method. The effect of path guidance on operators' operation is also evaluated.

2. Related work

2.1. Previous methods of surgical path planning

Surgical path planning of flexible needle in soft tissue has been studied in previous research works. The methods can be mainly divided into three categories [15]: numerical method, inverse solu-

tion method and search method. The numerical method consists of the probability density function method [16] and the objective function method [17]. Although numerical method is accurate in calculation, it will not be adaptable in complicated clinical environments with many kinds of irregular obstacles. Based on explicit geometric inverse kinematics, Duindam et al. [18] proposed a constant-time motion planning algorithm for steerable needles. This method depends on the idealized kinematics of the needle in a static and rigid environment. In addition, the inverse kinematic method sometimes cannot guarantee the solvability. The search method of the flexible needles mainly includes the artificial potential field method [19], the roadmap method [20], and the rapidly exploring random tree (RRT) method [21]. The search method is a relatively faster algorithm, but it cannot guarantee the optimal path. Nowadays, the numerical method, inverse kinematics method, and search methods are mostly used in geometrical environments and static environments.

Learning-based method was also applied for surgical path planning. Back et al. [22] proposed a path planning method for autonomous robotic gallbladder resection using probabilistic roadmap and reinforcement learning method. It was based on the precondition that the next state was predictable. For passive flexible tools, the use of reinforcement learning method was unpractical due to the unpredictable deformation of the tissue and tools. Schulman et al. [23], towards surgical suturing, proposed a non-rigid registration-based trajectory transfer for adapting a demonstrated tying trajectory from the training geometry to the testing ones.

2.2. Medical image processing using deep CNN

Deep CNN was recently widely applied for medical image processing, which demonstrated its capability for extracting and recognizing tissue's features from the medical image. Litjens et al. [24] reviewed the researches of deep learning in medical image processing. Nasr-Esfahani et al. [25] proposed a CNNs-based method for detecting vascular regions in digital subtraction angiography (DSA) images. Similarly, Wang et al. [11] proposed a novel deep learning-based framework for interactive medical image segmentation by incorporating CNNs into a bounding box and a scribble-based binary segmentation pipeline. Nie et al. [26] proposed a deep convolutional adversarial network framework for medical image synthesis. In our previous work [27], a CNNs-based method was proposed for automatic diagnosis of intracranial aneurysms in 3D Rotational Angiography based on a spatial information fusion method. The image understanding capability of deep CNN provides a promising way for surgical path planning by direct perception of medical images.

2.3. Deep learning-based path planning

Surgical path planning can be considered as a time sequence problem. Recurrent neural network (RNN) and LSTM are capable of extracting context-related features of time sequence data [28], such as natural language processing, pedestrian trajectory prediction, and so on. Fernando et al. [29] proposed an LSTM network architecture for pedestrian trajectory prediction, by using a soft-hardwired attention mechanism. Gupta et al. [30] introduced the LSTM network into GAN and proposed a Social GAN for pedestrian trajectory prediction, where the encoder and decoder are constructed by a set of LSTM networks. Each LSTM extracts features from a pedestrian's historical trajectory sequence. However, this method only takes pedestrian trajectory coordinates as network input, without considering the effect of environmental factors. Vinyals et al. [31] proposed a network structure that combines CNN and LSTM for image description task. Deep learning method, with the capability of spatial-temporal feature learning, provides

a promising solution for the challenges of surgical path planning. Our proposed surgical GAN differs mainly from Social GAN [30] and CNN-LSTM networks [31] by constructing the generator and discriminator of GAN with combined CNN-LSTM networks.

3. Materials and Methods

3.1. Problem Definition

3.1.1. Research goal

As shown in Fig. 1, the purpose of this work is to plan in real time a feasible future local path of guidewire tip in ES by estimating the current medical image (defined as surgical state image) and the historical trajectory of the guidewire tip.

The input surgical state image at time t is defined as X_{img}^t . The input historical trajectory is defined as X_{tr}^t . It consists of the guidewire tip positions at n times before the current time t . So, there is $X_{tr}^t = (X_p^{t-n+1}, X_p^{t-n+2}, \dots, X_p^t)$. $X_p^t = (x^t, y^t)$ is the guidewire tip position in image coordinate at time t . We denote the planned future local path as $Y \wedge_{path}^t = (Y \wedge_p^{t+1}, Y \wedge_p^{t+2}, \dots, Y \wedge_p^{t+n})$. $Y \wedge_p^{t+1} = (x \wedge^{t+1}, y \wedge^{t+1})$ is the guidewire tip position inside the planned local surgical path at time $t + 1$. The local path of ground truth collected from the surgeon’s demonstration is denoted as Y .

3.1.2. Assumptions and conditions

aaa

(1) As shown in Fig. 1 (a), for training sample collection and algorithm test, the simulated task is set to be delivering the guidewire tip from the aortic arch into the left common carotid artery (LCCA), left subclavian artery (LSCA), right common carotid artery (RCCA), or brachiocephalic trunk and so on. It is a necessary and relative difficult procedure in most of the intracranial endovascular surgeries including angiography diagnosis and treatment.

(2) The global aortic arch contours is known, which can be obtained by aortic arch angiography in actual ES. As shown in Fig. 1 (b), the starting point and the target point for the guidewire insertion tasks are manually set. The vascular center line

from the starting point to the target point is extracted. In each local image, the contours of the vascular and guidewire can be obtained in real time, and the vascular center line is used to represent the guidewire insertion task, as shown in Fig. 1 (c).

(3) The phrase of "path planning" in this work has some differences with typical path planning, especially for the idea of with no collisions. As shown in Fig. 1 (b) and (c), contacts between the guidewire and the vascular wall is even always necessary in many clinical cases, where the interventional radiologists delivers the guidewire tip pass the vascular bifurcation and into the target branch by making use of those contacts. Collisions between the planned guidewire path and the vascular contours is not absolutely unallowed in ES.

(4) Although the completeness of 3D anatomical images is better, 2D-DSA images with good real-time performance are still the gold standard for ES operation decision-making. To match current surgical patterns and habits, 2D images are used here as input for surgical path planning.

(5) Although fluoroscopy is a 2D technique, trajectory points of the guidewire tip in the 2D surgical state image can reflect both of the rotational and axial actions due to the guidewire tip is of J-shape. When the guidewire is rotated, pushed or pulled, the position of the guidewire tip will change. So, by taking historical trajectory points of the guidewire tip and the surgical state image as input, both of the rotational action and the axial action can be taken into account for trajectory planning.

(6) Due to more complex models are expensive and it is too early to test on actual tissues, manually designed vascular models (according to the medically used training model and human aortic arch anatomy) are used for training sample collection and algorithm test as detailedly shown in Section 3.3.

3.2. Proposed Framework

3.2.1. Image preprocessing

The proposed framework for guidewire path planning is shown in Fig. 2. The surgical state image at each time is preprocessed before input to the surgical GAN as follows:

1) First, in order to avoid the interference of other human tissues information in DSA image, the contours of the vessel and the guidewire are extracted using Qin’s VRBC-t-TNN vascular edge detection method [32], which is an effective method for accurately recovering vessel structures from the X-ray angiography images of moving organs or tissues. Although the vascular model used in this paper is rigid and has no motion and deformation, VRBC-t-TNN vascular edge detection method can be still used in further animal or in-human research, where the vessel is deformable. In addition, data collection experiments and user experiments are respectively conducted under camera and fluoroscopy in this work. Due to the adaptive thresholding method is adopted in Qin’s VRBC-t-TNN vascular edge detection algorithm [32], it is capable of preprocessing the camera images and fluoroscopy images.

2) Third, the extracted vascular contour is thinned into the vascular center line by using Zhang’s fast parallel thinning algorithm [33].

3) Fourth, the surgical environment near the guidewire tip plays the most important role during surgeons’ operation decision-making. So, the local surgical state image with a certain size is cropped by taking the guidewire tip as the center. The guidewire tip position is obtained in real-time by finding the endpoint of the thinned guidewire line. Relative to the whole surgical state image, the local surgical state image can be considered as an attention mechanism. Also, the sample diversity can be increased by using the local surgical state image [14].

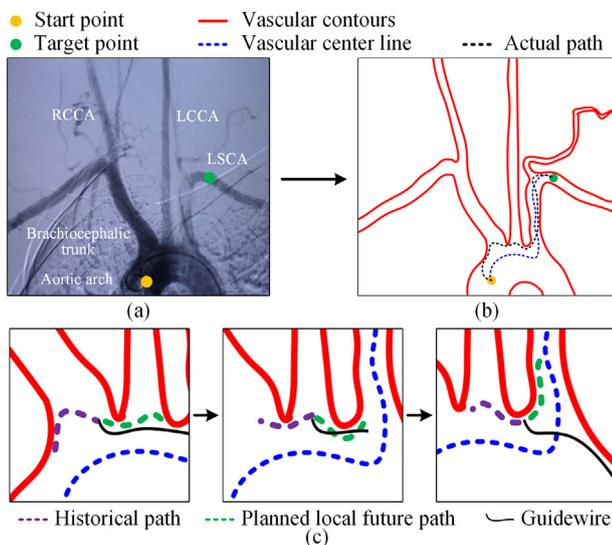


Fig. 1. Schematic of guidewire-tip local path planning: (a) an aortic arch angiography with manually set start and target points of the guidewire insertion task, (b) extracted vascular contours, center line from the start point to the target point and final actual path of guidewire insertion, (c) local images with historical path and planned local future path.

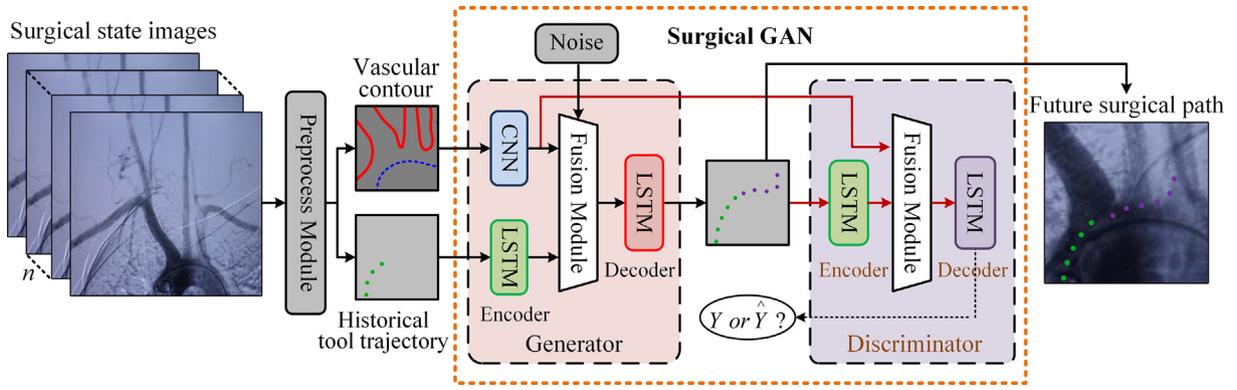


Fig. 2. Architecture of the proposed framework for guidewire path planning. In Preprocess Module, vessel and guidewire contours, as well as the historical tool trajectory, are extracted from the surgical state image sequence (DSA images). Then, the preprocessed information is input to surgical GAN. A new Generator is the main innovation of surgical GAN, which can extract and fuse the spatial and temporal features in the unstructured surgical environment. It is enabled by combining CNN and LSTM within the encoder, which is followed by a Fusion Module. The Decoder generates future local surgical path in real time.

The preprocessed local surgical state image and the historical guidewire tip positions sequence are then input to surgical GAN. Surgical GAN extracts the spatial features of medical image and temporal features of historical guidewire trajectory, and then predicts the future surgical path.

3.2.2. Surgical GAN

Conditional Generative Adversarial Networks (CGAN) was introduced as an extension of GAN [34]. GAN has demonstrated that it can produce realistic samples from randomly distributed input based on adversarial training. However, in an unconditional generative model, there is no control on the models of the data being generated. It means that we cannot obtain desired data corresponding to certain conditions. By conditioning the model on extra information η . CGAN is effective to direct the data generation process. η could be any kind of auxiliary information, such as class labels or data from other modalities.

Like GAN, CGAN also consists of two adversarial models: a Generator G takes the prior noise $p_z(z)$ and η as input and outputs a generated sample $G(z|\eta)$. A Discriminator D takes a generated sample or a real one as input x and outputs a single scalar $D(x|\eta)$. $D(x|\eta)$ represents the possibility that x is a real sample rather than a generated one. The Generator and Discriminator play the min-max game with a value function $V(D, G)$ and can finally reach equilibrium [34], where both Generator and Discriminator achieve the optimal capability.

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x|\eta)] + E_{z \sim p_z(z)} [\log (1 - D(G(z|\eta)))] \quad (1)$$

Our work is primarily inspired by the recent Social GAN framework for pedestrian trajectory prediction [30], which is essentially a kind of CGAN. The historical trajectory of each person in a scene is taken as conditional information. It was demonstrated that Social GAN can successfully generate socially acceptable trajectories for every pedestrian in the scene. Surrounding state information was not considered in Social GAN, because LSTM was not capable of modeling 2D image data. However, surgical state information in the medical image is crucial for surgical path planning. To solve this problem, we propose surgical GAN as shown in Fig. 2, which combines CNN and LSTM to construct the Generator and Discriminator. The Generator used the surgical state image and historical tool trajectory as the conditional information and generate the future tool path as real as possible. The Discriminator

is designed to distinguish the generated tool path from a real one in the training sample. The surgical GAN will be illustrated in detail as follows.

3.2.3. Generator of Surgical GAN

The Generator of surgical GAN adopts encoder-decoder structure as shown in Fig. 3. The encoder is used for encoding the surgical state image and historical tool trajectory into intermediate features. Then, the intermediate features are decoded by the decoder to generate the future surgical path of the guidewire tip. In order to combine the feature outputs of the CNN and LSTM, a fusion module is introduced between the encoder and decoder.

3.2.4. Encoder of the Generator

First, the surgeon makes decisions on the surgical path mainly according to the morphological characteristics and spatial relationship of vascular contour and the guidewire. So, CNN is applied to extract the features in the surgical state image. The l th convolutional layer can be given by

$$h_l^t = \sigma(h_{l-1}^t W_l + b_l) \quad (2)$$

where h_l^t is the output feature map of the l th convolutional layer and h_{l-1}^t is the output of the previous layer, $h_{l-1}^t|_{l=1} = X_{pr_img}^t$, W_l represents the kernels of the l th convolutional layer, b_l is the bias term; $\sigma(\cdot)$ is the nonlinear activation function, and ReLU is used.

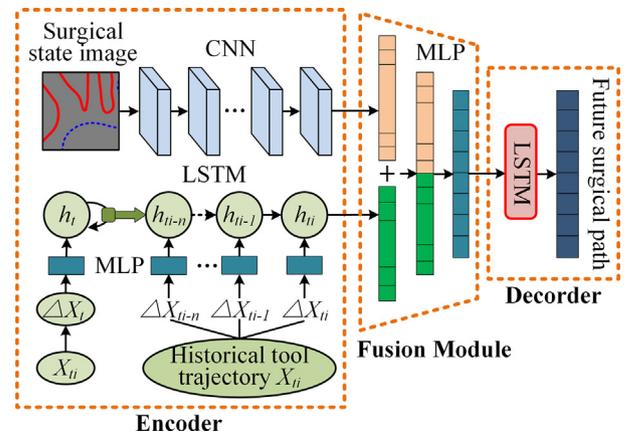


Fig. 3. Architecture of the generator of surgical GAN. A MLP is used to fuse the extracted spatial and features from surgical state image and historical tool trajectory.

The final feature output of the CNN is the output of the last fully connected layer as given in

$$h_f^t = \sigma(W_f h_{f-1}^t + b_f) \quad (3)$$

where h_f^t and h_{f-1}^t are respectively the output feature vector and input feature vector of the f th fully connected layer, W_f is the weight matrix, b_f is the bias term.

Second, surgical path planning can be considered as a sequential problem. The surgeon adopts certain strategies to deliver the guidewire tip for a given surgical task. The surgical path of the guidewire tip can be regarded as an explicit representation of the surgical strategy with context. In this paper, we use LSTM to identify the sequential character of the tool's historical trajectory and then predict its future surgical path. It should be noted that multi-LSTMs could also be used when planning surgical paths for more than one tools.

Third, rotation-and-translation invariance of the sequential character of the guidewire tip trajectory needs to be considered. To increase sample diversity, data augmentation is performed by rotating and translating the images. It is obvious that the spatial relationship between the trajectory and vascular contour is independent of the image's rotation and translation. On the contrary, the absolute positions of trajectory points change in such conditions, which might cause gradient disappearance and explosion during network training. So, we normalize the absolute coordinates of the trajectory points by calculating the position difference between adjacent points. The sequence of the position differences of n trajectory points at time t is defined as ΔX_{tr}^t , which can be expressed as:

$$\Delta X_{tr}^t = (\Delta X_{tr}^{t-1}, \Delta X_{tr}^{t-2}, \dots, \Delta X_{tr}^{t-n}) \quad (4)$$

$$\Delta X_{tr}^{t-i} = (x^{t-i} - x^{t-i-1}, y^{t-i} - y^{t-i-1}) \quad (5)$$

It should be noted that the position difference of trajectory points only consists of two numerical values, i.e. difference in horizontal and vertical coordinates. For the network weights balance, a single layer MLP is used for embedding the position differences into a high-dimensional vector e_{tr-e}^{t-i} as given in:

$$e_{tr-e}^{t-i} = \phi_1(\Delta X_{tr}^{t-i}, \Delta y_{tr}^{t-i}, W_{emb-G}) \quad (6)$$

where $\phi_1(\cdot)$ represents the single layer MLP using ReLU as the activation function, W_{emb-G} is the weight of the MLP.

Then, the sequence of embedded vectors of n trajectory points e_{tr}^t is used as the input of encoder LSTM of the Generator as expressed in:

$$o_{eG}^{t-i} = LSTM_{eG}(h_{lstm-e}^{t-i-1}, e_{tr}^{t-i-1}, C_{lstm-e}^{t-i-1}, W_{eG}) \quad (7)$$

where o_{eG}^{t-i} , h_{lstm-e}^{t-i-1} , C_{lstm-e}^{t-i-1} , W_{eG} are respectively the output feature vector, hidden state, cell state and weights of encoder LSTM of the Generator at time $t-i$.

3.2.5. Fusion module

In order to combine the extracted features of the surgical state image and historical guidewire tip trajectory, we introduce a fusion module between the encoder and the decoder. The output feature map of the CNN is flattened into a one-dimensional vector, and then concatenated with the output feature vector of decoder LSTM of the Generator. Then, the concatenated vector is input to an MLP to non-linearly fuse with the features, as given in

$$e_{fus-G}^t = \phi_2\left(\left(h_f^t, o_{eG}^t\right), W_{fus-G}\right) \quad (8)$$

where $\phi_2(\cdot)$ represents the MLP in fusion module using ReLU as the activation function, W_{fus-G} is the weight of the MLP in fusion module of the Generator.

3.2.6. Decoder of the Generator

The aim of the decoder is to generate the future surgical path by taking the fused feature as input. LSTM is used to construct the decoder, and the position difference $\Delta Y \wedge_{path}^{t+j}$ between future surgical path points at time $t+j$ can be calculated by

$$\Delta Y \wedge_{path}^{t+j} = LSTM_{dG}(h_{lstm-dG}^{t+j-1}, e_{fus}^{t+j-1}, C_{lstm-dG}^{t+j-1}, W_{dG}) \quad (9)$$

where $h_{lstm-dG}^{t+j-1}$, $C_{lstm-dG}^{t+j-1}$, W_{dG} are respectively the hidden state, cell state and weights of decoder LSTM at time $t+j-1$.

The guidewire tip position $Y \wedge_{path}^{t+j}$ in the future surgical path at each time step can be obtained by adding iteratively the generated position differences sequence $\Delta Y \wedge_{t+j}$ to the current guidewire tip position, as given in

$$Y \wedge_{path}^{t+j} = Y \wedge_{path}^{t+j-1} + \Delta Y \wedge_{path}^{t+j} \quad (10)$$

3.2.7. Discriminator of surgical GAN

The Discriminator is also constructed by encoder-decoder structure. LSTM is used to construct the Discriminator. It takes the position difference sequence of a planned surgical path or a sample as input. Kevin et al. [35] demonstrated that weights sharing between the Generator and Discriminator is beneficial in improving the training efficiency. So, the encoder LSTM of the Generator shares weights with the encoder LSTM of the Generator. The output feature is then combined with the feature output of CNN in Generator by a fusion module. The feature output of CNN from surgical state image plays the role of additional information η of the Discriminator. LSTM is then used as a decoder. It takes the combined feature as input and outputs the possibility of a surgical path is a real sample or generated one. It should be pointed out that the decoder LSTM in Discriminator and Generator do not share weights. In addition, the position difference at each time is also embedded into a high-dimensional vector e_{tr-dsc}^{t+j} using a single layer MLP.

The embedded vector e_{tr-dsc}^{t+j} is then input into the encoder LSTM of the Discriminator as given by:

$$h_{eD}^{t+j} = LSTM_{eD}(h_{eD}^{t+j-1}, e_{tr-dsc}^{t+j-1}, C_{eD}^{t+j-1}, W_{eD}) \quad (11)$$

where h_{eD}^{t+j-1} , C_{eD}^{t+j-1} , W_{eD} are respectively the hidden state, cell state and weights of the decoder LSTM of the Discriminator at time $t+j-1$.

Then, a fusion module is used for embedding the output feature map of the Generator CNN and output feature vector of the encoder LSTM in Discriminator, which can be expressed as:

$$e_{fus-D}^{t+j} = \phi_4\left(\left(h_f^t, h_{eD}^{t+j}\right), W_{fus-D}\right) \quad (12)$$

where $\phi_4(\cdot)$ represents the MLP in fusion module using ReLU as the activation function, W_{fus-D} is the weight of the MLP in fusion module of the Discriminator.

Then, the output of fusion module is input to the decoder LSTM of Discriminator, which can be expressed by:

$$h_{dD}^{t+j} = LSTM_{dD}(h_{dD}^{t+j-1}, e_{fus-D}^{t+j-1}, C_{dD}^{t+j-1}, W_{dD}) \quad (13)$$

$$o_{dsc}^t = h_{dD}^{t+n} \quad (14)$$

where h_{dD}^{t+j-1} , C_{dD}^{t+j-1} , W_{dD} are respectively the hidden state, cell state and weights of the decoder LSTM of the Discriminator at time $t+j-1$, o_{dsc}^t is the final output of the Discriminator at time $t+n$ ($n=8$).

It should be noted that the Discriminator only works during network training but not path planning procedure. During path planning application, the final output of surgical GAN is calculated by iteratively adding the position differences generated by the Generator to the current guidewire tip position.

3.3. Data Collection and Network Training

Due to ethical limitations, it is impractical to collect large-scale training data from clinics. So, we develop a laboratory platform for training data collection as shown in Fig. A.1 and A.2. The main components of the laboratory platform for training data collection include the following: designed vascular models and a loach guidewire (RF*GA35153M, Terumo Corporation, JP, with a diameter of 0.9 mm, length of 150 mm, and a J-shape tip of 3 mm length) used for simulated guidewire insertion, an industrial camera (with 1280*960 pixels, 15 fps, a wide-angle lens of 145° FOV and 2.8 mm focal length) used for capturing the image of the transparent vascular model and the guidewire, a set of LED and light shield to provide a stable photoenvironment. Considering sample diversity, a sufficient number of vascular models with different shapes are needed for data collection. With minor adjustments according to the actual anatomic structure of human vessels, we design 10 vascular models with different vascular radius, bending degrees, including angles and relative positions between vascular bifurcations, as shown in Fig. A.1. The set angles, diameters etc are realistic and within the normal range of human vascular anatomical morphology parameters. It should be noted that because deformable vascular models are costly, rigid vascular models are used in this work for preliminary verification of the proposed method.

The simulated operation task is set to be delivering the guidewire tip from the aortic arch into the left common carotid artery (LCCA), left subclavian artery (LSCA), right common carotid artery (RCCA), or brachiocephalic trunk and so on. On the one hand, the given task is necessary for most of the intracranial endovascular surgeries of angiography diagnosis and treatment. During intracranial endovascular surgeries, the path from aortic arch to those three branches is the only way to get to the target in common carotid artery, vertebral artery, and further branches. On the other hand, in clinical practice, the given task is a relatively difficult and time-consuming stage in the whole process of delivering the guidewire tip from the femoral artery to the intracranial nidus.

Ten surgeons, aged 30–50, are invited to implement the tasks using a medical guidewire (RF*GA35153M, Terumo Corporation, JP) for sample collection. The simulated operation in each vascular branch of each model is repeated 5 times by each surgeon. If the surgeon missed the target, he/she was asked to repeat the action in order to exclude wrong operation and sample data. There are 4 to 6 branches that can be effectively inserted in each of the 10 vascular models. One simulated operation attempt is considered as a demonstration. The image of the vessels and the guidewire at every time of the demonstrations is recorded by the camera, which is defined as the surgical state image. The surgical environment near the guidewire tip is important for surgeons' operation decision-making, so the local surgical state image with a certain size is cropped from the whole surgical state image by taking the guidewire tip as the center. Each sample consists of a preprocessed local surgical state image and a segment of historical guidewire tip trajectory as the network input, as well as a segment of the future surgical path as the label. Finally, 2.5×10^3 demonstrations (5 demonstrations/branch/surgeon \times 5 branches/model \times 10 models \times 10 surgeons) are collected and the sample datasets containing 3.5×10^5 samples (140 samples/demonstration \times 2.5×10^3 demonstrations) are constructed.

We implement the proposed surgical GAN based on Pytorch and a workstation (with an NVIDIA GTX 1080Ti GPU). 3.4×10^5 collected samples are used as the training set, while 1×10^4 samples are used as the validation set. 1×10^4 samples collected with an endovascular evaluator (EVE) is used as the test set. Adam optimizer is used for optimization. The batch-size, Discriminator learning rate and Generator learning rate are respectively 64, 5×10^{-5} and 8×10^{-5} . During tests, the calculating speed is about 20 fps, which meets the requirement of real-time path planning in operation.

4. Experiments and Results

In this section, experiments are conducted to evaluate the capability of the proposed surgical GAN framework for guidewire path planning in ES. Firstly, comparative tests are conducted among the fully surgical GAN, several baseline networks and nonrigid registration method [23]. Secondly, user experiments are conducted to evaluate the guidance effect of the planned surgical path on surgeon's operation. All metrics are assessed using the non-parametric Kruskal–Wallis H significance test in IBM SPSS Statistics 24 (a value of $p < 0.05$ is considered statistically significant).

4.1. Effect of CNN achitecture and path planning length

On one hand, different CNN architectures can theoretically be used in the proposed surgical GAN framework. So, the effects of different CNN on path planning performance should be tested. Using the same training and test data, path planning tests are conducted by integrating different typical CNNs including VGG13, DenseNet121, Inception V3, and ResNet50.

On the other hand, since context features of historical tool path are fused for path planning, the influence of path planning length n on path planning performance needs to be investigated. Experiments are conducted by setting n from 4 to 16. It should be noted that path planning length is equal to historical tool path length, because the encoder LSTM of Generator share weights with that of Discriminator.

For quantitative evaluation, six evaluating metrics are defined as follows. The average diameter of the vessel for one guidewire insertion demonstration dia_{vessel} is defined as the twice of the average distance between the points in the vessel central line and the closest points in the vessel contour line, which can be given as:

$$dia_{vessel} = \sum_1^m 2dis_{cent-cont} / m \quad (15)$$

where $dis_{cent-cont}$ is the distance between a point in the vessel central line and the closest point in the vascular contour line, m is the number of the points in the vessel center line from the start point to the target point in one guidewire insertion demonstration.

(1) *Average Deviation (AD)*: it donates the average deviation between the points of the real path and the planned path. It reflects the overall deviation. Then, the *Relative Average Deviation (RAD)* is defined as the ratio of AD and dia_{vessel} .

(2) *First Point Deviation (FPD)*: it is the deviation between the first point in the real path and that in the planned path. For dynamic path planning, the first path point nearest to the guidewire tip plays an important role. The *Relative First Point Deviation (RFPD)* is defined as the ratio of FPD and dia_{vessel} .

(3) *Last Point Deviation (LPD)*: it defines the deviation between the last point of the real path and that of the planned path. It reflects the final accuracy of the prediction. The *Relative Last Point Deviation (RLPD)* is defined as the ratio of LPD and dia_{vessel} .

The results are shown in Table 1. It can be seen that AD, RAD, FPD, RFPD, LPD and RLPD are almost at the same level despite different CNN architectures. It indicates that, the CNN within the proposed surgical GAN can all extract the features in the surgical state images. In addition, LPD and RLPD get the minimum value of 1.22 mm and 5.96% with DensNet121. AD, RAD, FPD and RFPD achieve the minimum values of 1.16 mm, 5.67%, 0.48 mm and 2.35% with Inception V3. So, Inception V3 is selected for the subsequent experiments.

The results of tests with different path planning lengths are shown in Table 2. It can be found that FPD and RFPD reduce rapidly from the maximum value of 0.55 mm and 2.69% to 0.40 mm and 1.95% when n increases from 4 to 8. Then, FPD and RFPD convergence slowly to the minimum value of 0.39 mm and 1.91% when n gets to 12. The reason might be that when n is 12, the network obtains the most context information of historical tool path. But, when n gets to 16, the part of context information that is too far from current time offers little help for future path planning. LPD and RLPD achieve the minimum values when n is 8. The cause might be that the remote predicted path point has little relevance with the current position, although more context information is beneficial for future path planning. AD and RAD also achieve the minimum value when n is 8. So, n is set to be 8 for the subsequent experiments.

4.2. Methods Comparison

Social GAN [30] with a single LSTM is defined here as GAN-LSTM. A network composed of inceptionV3 and LSTM is trained via supervised learning but not adversarial learning. It is defined as Inception V3-LTSM [31]. GAN-LSTM and InceptionV3-LTSM are taken as the baseline networks. The effect of surgical environment information and adversarial learning on surgical path planning can be analyzed via comparison between fully surgical GAN, GAN-LSTM and InceptionV3-LTSM. The effect of surgical environment information is further evaluated by comparison between fully surgical GAN and the non-rigid registration method [23]. The non-rigid registration method is implemented using Matlab with the collected training sample data in this work.

The quantitative comparison results are shown in Table 3. First, it can be seen that AD, RAD, FPD, RFPD, LPD and RLPD of full surgi-

Table 1
Effect of different CNN architectures.

Metrics	Surgical GAN_VGG	Surgical GAN_DenseNet	Surgical GAN_Inception	Surgical GAN_ResNet
AD(mm)	1.23	1.18	1.16	1.20
RAD	7.24%	6.95%	6.83%	7.07%
FPD(mm)	0.49	0.50	0.48	0.52
RFPD	2.89%	2.94%	2.83%	3.06%
LPD(mm)	1.35	1.22	1.24	1.48
RLPD	7.95%	7.18%	7.30%	8.71%

* Path planning length n is 6 in the tests with different CNN architectures.

Table 2
Effect of different path planning lengths.

Metrics	Surgical GAN_4	Surgical GAN_6	Surgical GAN_8	Surgical GAN_10	Surgical GAN_12	Surgical GAN_16
AD(mm)	1.27	1.16	1.02	1.12	1.32	1.64
RAD	7.48%	6.83%	6.01%	6.59%	7.77%	9.66%
FPD(mm)	0.55	0.48	0.40	0.45	0.39	0.46
RFPD	3.24%	2.83%	2.36%	2.65%	2.30%	2.71%
LPD(mm)	1.65	1.53	1.24	1.60	1.88	2.57
RLPD	9.72%	9.01%	7.30%	9.42%	11.07%	15.13%

* Inception V3 are selected in these tests. Surgical GAN _{n} denotes the network with path planning length of n .

cal GAN are obviously lower (up to 59.49%) than those of GAN-LSTM. It indicates that the surgical state image features are successfully recognized. Because of the lack of environmental information, GAN-LSTM shows lower accuracy of path planning. AD, RAD, FPD, RFPD, LPD and RLPD of full surgical GAN are overall lower than those of Inception V3-LSTM by up to 69.78%. It indicates that the generative adversarial training method is beneficial for improving the accuracy of path planning. Second, because the non-rigid registration method just generates a fixed whole path from start point to the target, only AD and RAD are meaningful for comparison. AD and RAD of full surgical GAN are 72.94% lower than those of the non-rigid registration method. Third, the size of the surgical state image is 200×200 pixels and 45×45 mm. AD, RAD, FPD, RFPD, LPD and RLPD of the proposed surgical GAN are respectively 1.02 mm, 4.99%, 0.40 mm, 1.95%, 1.24 mm and 6.06%. The path planning accuracy is at the same scale of the operating accuracy of our previous ES robot (with average error of 0.18 mm and maximal error of 1.4 mm), which has successfully assisted a interventional radiologist to implement bilateral carotid angiographies in human[36]. The operation task for demonstration and algorithm test in this work, delivering the guidewire tip from the aortic arch into the left common carotid artery (LCCA), left subclavian artery (LSCA), right common carotid artery (RCCA), or brachiocephalic trunk and so on, is a necessary and difficult procedure during bilateral carotid angiographies. It indicates that the proposed surgical GAN path planning method can be used for operating guidance.

The results of guidewire path planning with surgical GAN in two vascular branches are shown in Fig. 4. It should be noted that the ground truth path is here defined as a recorded guidewire tip path in a successful demonstration of an expert surgeon. The reason is that there is currently no model that can calculate an absolutely optimal path as the ground truth path. This research work just aims to develop a model that can plan an effective path by learning the features contained in experienced surgeons' demonstration data. The skills of experienced surgeons are considered to be better than that of novice surgeons, which can be partly represented by the operation path. So, to evaluate the proposed path planning method, it is a rational and currently good way by taking the demonstration of the expert surgeon as the ground truth path. It can be seen that the planned path matches well with the ground truth, although there is a certain deviation between them. The reason for the deviation is partly that the variance of the path point distribution between different training samples. Guidewire's deformation caused by operation resistance results in a variation of its forward velocity. This variation leads to small or large position differences between planned path points under certain probability. In addition, this deviation is larger near the vascular branches. The reason might be that the shape complexity of the vascular branches decreases the feature recognition accuracy of the network. But the deviation can reduce again after the guidewire tip passes through the branches as shown in Fig. 4.

For comparing analysis between the proposed method and the non-rigid registration method, non-rigid registration is used to

Table 3
Results of Method Comparison.

Metrics	GAN-LSTM[30]	Inception V3-LSTM[31]	Full surgical GAN	Non-rigid registration[23]
AD(mm)/ RAD	2.07/ 12.19% (55.60% ↓*)	2.49/ 14.66% (63.09% ↓)	1.02/ 6.01%	3.77 / 22.20% (72.94% ↓)
FPD(mm)/ RFPD	0.84/ 4.95% (52.15% ↓)	0.74/ 4.36% (46.20% ↓)	0.40/ 2.36%	—
LPD(mm)/ RLPD	3.06/ 18.02% (59.49% ↓)	4.10/ 24.14% (69.78% ↓)	1.24/ 7.30%	—

* ↓ indicates that the metric value of Full surgical GAN is lower than that of other methods.

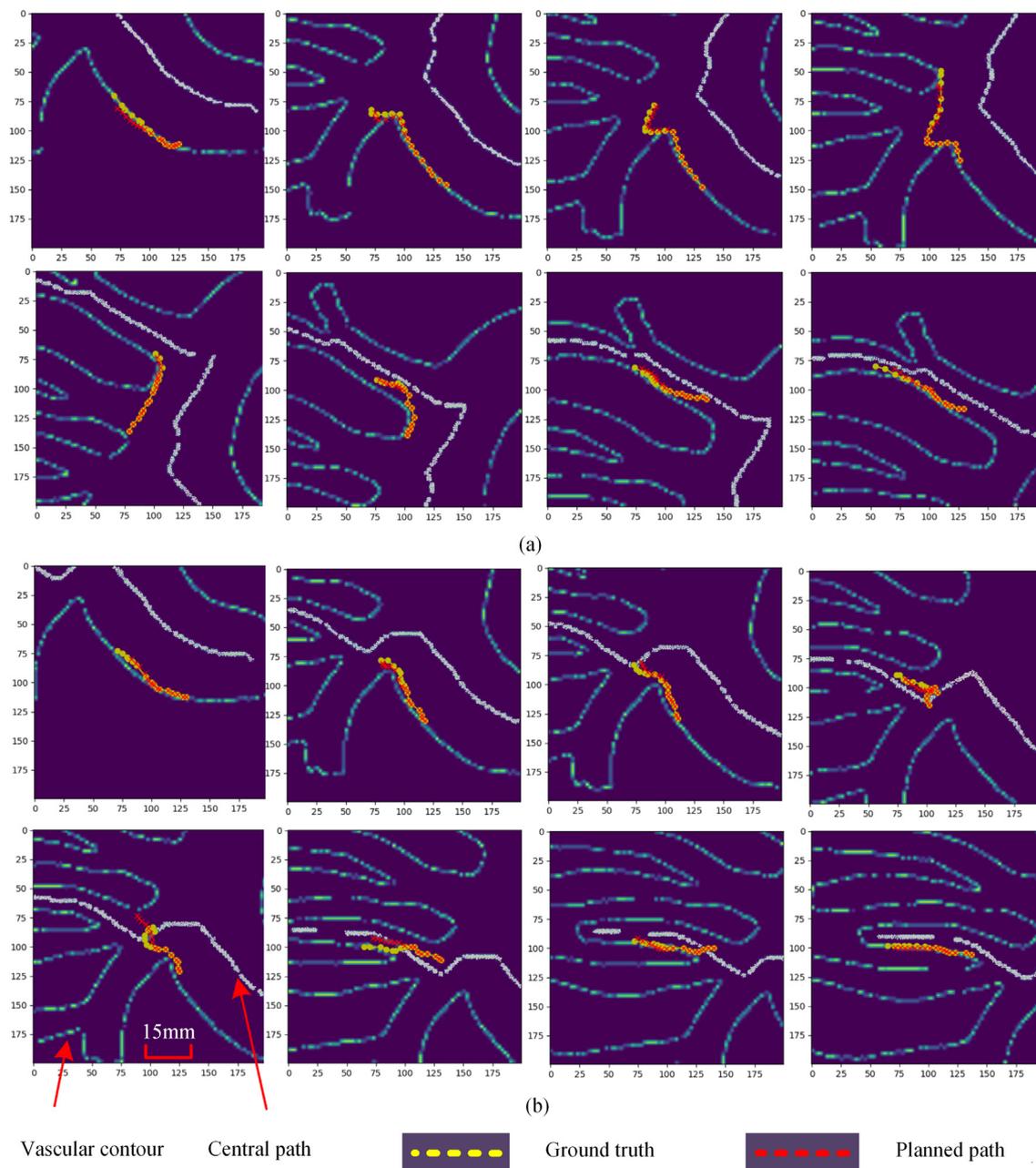


Fig. 4. Results of surgical path planning with surgical GAN in: (a) vascular branch I, (b) vascular branch II. The ground truth (a recorded guidewire tip path in a successful demonstration from an expert surgeon) are marked in yellow. The planned path are marked in red.

obtain a warping function to map guidewire tip paths from a sample vascular model into an other anatomically similar but shape/scale/orientation different target vascular model. The results of guidewire path planning with the non-rigid registration method for two vascular branches in the target model are shown in Fig. 5. The legend for the red line: "ground truth in the sample vas-

cular model" means a recorded guidewire tip path within the sample vascular model in a successful demonstration of an expert surgeon. The legend for the blue line: "ground truth in target vascular model" denotes the recorded guidewire tip path within the target vascular model in a successful demonstration of an expert surgeon. It can be seen that the overall trend of the planned path

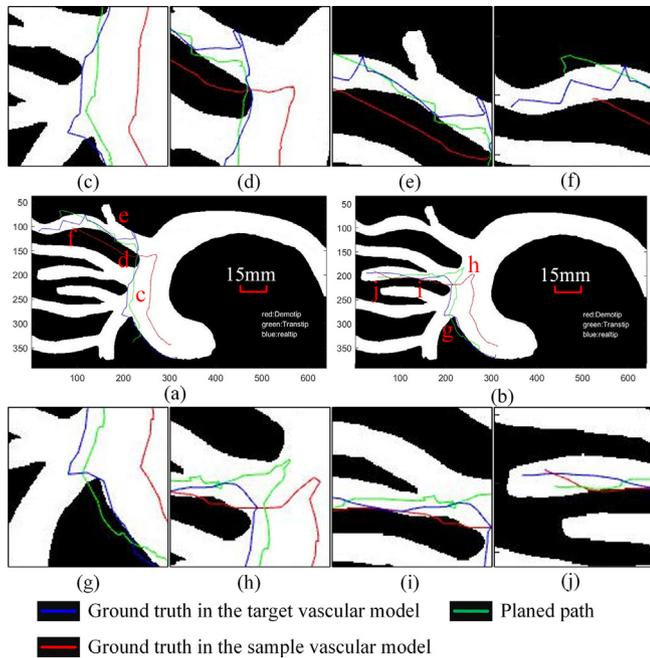


Fig. 5. Results of surgical path planning with the non-rigid registration method in: (a) vascular branch I, (c) to (f) are large versions of several different local images at points c to d in (a), (b) vascular branch II, (g) to (j) are large versions of several different local images at points g to j in (b).

and the ground truth in the target vascular model are similar. But the deviation between the planned path and the ground truth in the target vascular model is relatively large. As shown in Fig. 5 (h), the planned path even shows a similar trend with the ground truth in the sample vascular model but a different trend with the ground truth. The planned path in Fig. 5 (d), (f) and (g) are partially out of the vascular contour. Comparison between Fig. 4 and Fig. 3 shows that the planned path with the proposed surgical GAN matches better with the expert surgeon’s operating path than that with the non-rigid registration method.

4.3. User experiment

In this section, we further evaluate the guidance effect of the planned path on the surgeon’s operation through objective and subjective metrics. We develop a simulated clinical scenario in an operating room as shown in Fig. 5. The surgeon remotely manipulates the guidewire via a master-follower ES robot. It should be noted that The follower robot does not automatically navigate to the target, but manipulates the guidewire under the control of the surgeon (via the master controller). During the experiments, the surgeon manipulates the operating handle of the master controller. The sensing device of the master controller detects the axial and rotational motion of the operating handle, which represents the operating action of the surgeon. At the same time, the detected action information is sent to the controller of the follower robot. By accurately reproducing the surgeons action, the follower robot manipulates the guidewire to deliver the guidewire tip towards the target point. An endovascular evaluator (EVE) is used to provide a relatively realistic surgical environment. It is remolded according to the actual vascular anatomy from a woman’s 3D CT data and made of silicon by 3D printing process. It means that the anatomy of the silicon phantom is realistic. In addition, such a silicon phantom has been used as an endovascular evaluator for skills training of interventional radiologist. So, it is practical

and reasonable to use the silicon phantom for preliminary verification of the proposed method in this work. With contrast agent injection, the contour of the vessels is distinct from the X-ray image before subtraction, as shown in Fig. 5 (c). Without contrast agent injection, the contour of the guidewire is distinct from the X-ray image after subtraction, as shown in Fig. 5 (d). At each time step, the preprocessed image is used as input of the surgical GAN. We visualize the planned surgical path and match it with the surgical state image, which is used as the interface to the operator (a GIF is given in the Supplementary material as GIF. A.2 to show the path guidance procedure).

Three novice female and seven novice male participants, ages 25–35, are invited to implement the guidewire insertion task. The operators are asked to perform guidewire insertion tasks similar to those for sample collection in Section 2. The tasks are performed in two different modes: guidance and free. Each participant performs the task for 5 times in free mode, where no path guidance is provided. After ten days, the participants are considered to be unfamiliar again to the vascular model. Then, each participant performs the task for 5 times in guidance mode, where path guidance is provided. Fig. 6.

The qualitative results are shown in Fig. 7. According to the planned path in Fig. 7 (a) and (b), the guidewire tip should be pulled back in the aortic arch towards LSCA, which is reasonable. As shown in Fig. 7 (b), the smoothness of the planned path is decreased. It might be caused by the complex shape near the entrance of the Innominate artery. According to the planned path in Fig. 7 (c) and (d), the guidewire tip should be rotated and pushed forward. It is rational because the guidewire tip is near the entrance of LSCA but its orientation deviates from the entrance. From Fig. 7 (e) and (f), the guidewire tip is pulled by an overlarge distance and passes the entrance of LSCA. Then, as shown in Fig. 7 (g) and (h), the operator tries again and pushes the guidewire tip towards the entrance of LSCA. During the procedure from Fig. 7 (e) to (h), the planned path remains oriented to LSCA. It provides the right guidance to the operator. In Fig. 7 (i) to (l), the guidewire tip is inserted into LSCA and gradually gets to the target. The results indicate that the trained surgical GAN has the ability of planning the acceptable local surgical path in real time.

For quantitatively evaluating the effect of path guidance on surgeons’ operation, following objective metrics are defined:

- (1) OD(s): Operation duration for task completion. The extension of OD might aggravate X-ray radiation damage to both the patient and surgeon [37].
- (2) TD(m): Travel distance of the guidewire tip. Larger TD indicates more touch between guidewire and the vessel and then increased risk of vascular damage.
- (3) OA: Operation attempts for task completion. The surgeon always attempts several times to pass the guidewire tip through some complex vascular region. Larger OA also increases the risk of vascular damage.
- [4] TT(°): Trajectory tortuosity of the guidewire tip, which represents the smoothness of the guidewire insertion. Smaller TT represents a better operation quality. The trajectory tortuosity near point p_i can be represented by the including angle η_i between line $p_{i-1}p_i$ and line $p_i p_{i+1}$. η_i can be calculated by cosine law. Then, TT can be considered as the average value of η_i .

The objective evaluation results are shown in Table 4 and Fig. 8. It can be seen that OD, TD, OA and TT in guidance mode are improved compared with those in free mode. In addition, the p value for the four metrics are all smaller than 0.05. It indicates that the effect of path guidance on the operators’ operation is significant. As shown in Fig. 8, the smoothness of the guidewire

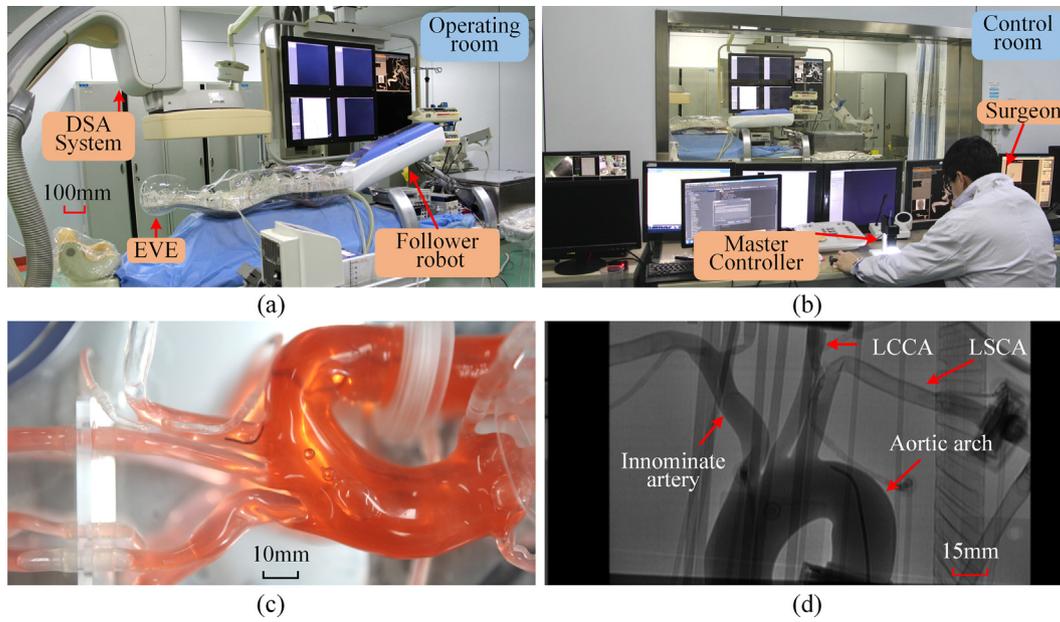


Fig. 6. Experimental platform in an operation room: (a) EVE and follower robot, (b) A surgeon manipulating the master controller, (c) Enlarged view of EVE, and (d) Vascular contour in the X-ray image before subtraction.

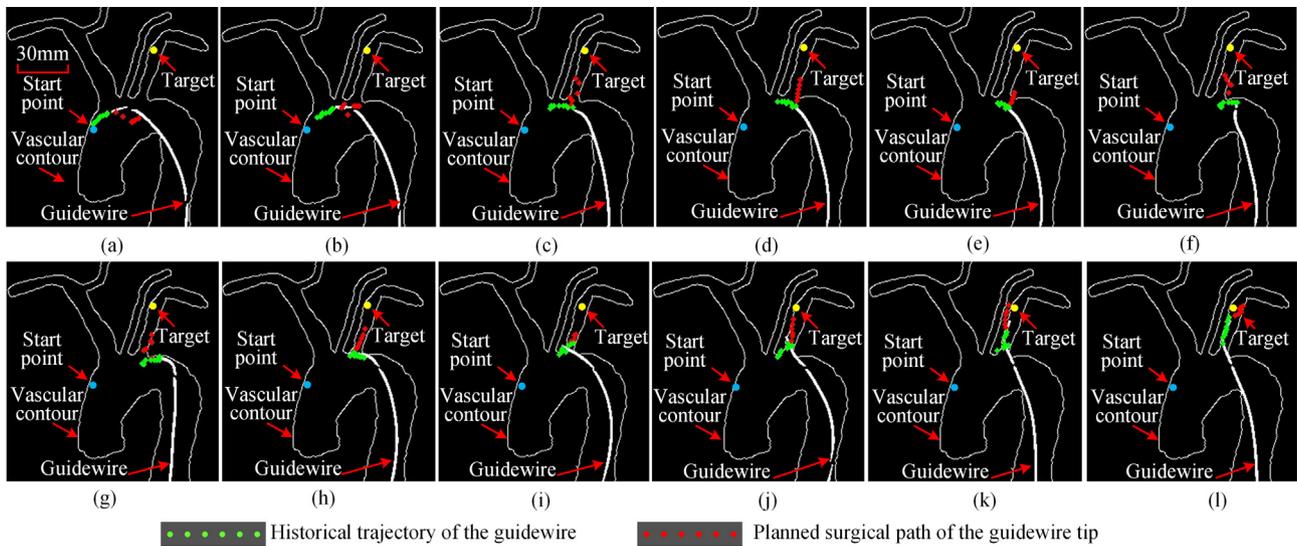


Fig. 7. Guidewire tip path planning in a user experiment. As the guidewire tip moves from the start point in the aortic arch to the target in LSCA, the algorithm always plans reasonable surgical path. It can provide visual guidance for novice surgeons, who are not familiar with the operation.

Table 4
The Quantitative Results of the User Experiment.

Metrics	Operation mode	Operator number										p value
		No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	
OD(s)	Free	51.2	66.4	53.6	60.1	52.2	62.9	57.7	64.8	51.1	54.3	0.006 <0.05
	Guidance	43.8	46.3	40.2	52.4	47.6	52.2	53.1	56.4	43.7	43.9	
TD(m)	Free	89.8	105.6	108.3	126.1	90.3	119.3	109.1	122.5	88.4	90.2	0.002 <0.05
	Guidance	74.8	80.6	88.4	79.2	73.9	76.2	79.8	77.2	76.1	75.4	
OA	Free	2	3	3	4	2	4	3	4	2	2	0.005 <0.05
	Guidance	1	1	2	1	1	1	2	2	1	1	
TT(°)	Free	161.3	161.7	161.4	160.8	162.4	163.1	162.6	163.6	162.8	161.6	0.002 <0.05
	Guidance	168.1	163.9	166.3	162.5	167.7	163.3	161.9	164.1	164.9	166.4	

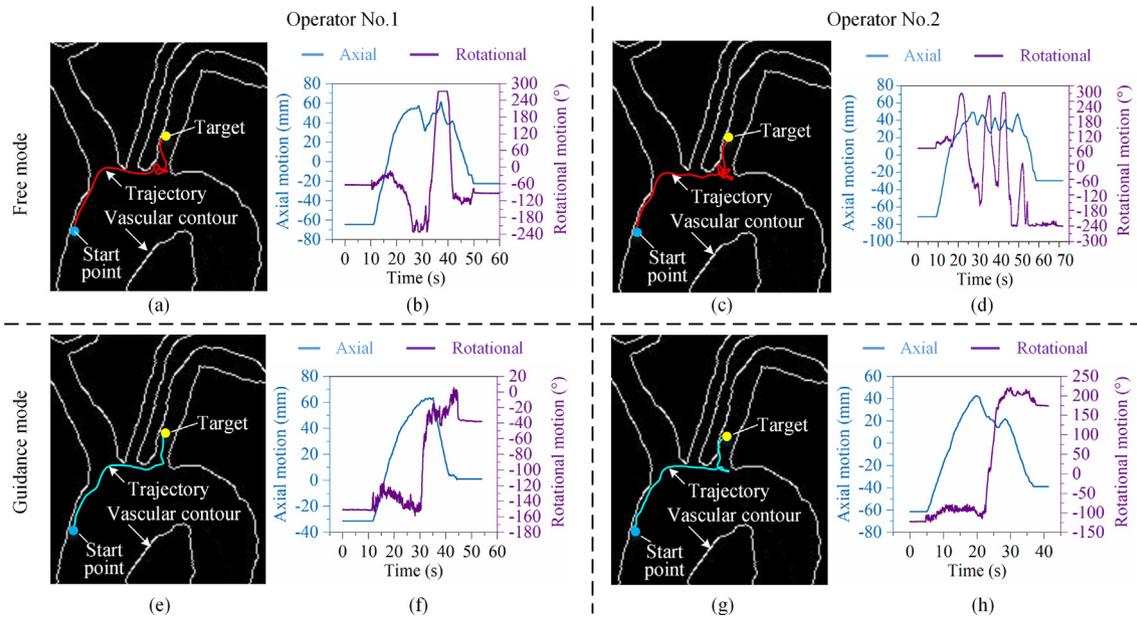


Fig. 8. Results between guidance mode and free mode of two operators' tests. (a) (c) (e) (g) are the guidewire tip trajectories, (b) (d) (f) (h) are operators action detected by the maser controller. The axial motion represents the push (increase) and pull (decrease) action. The rotational motion represents the clockwise (increase) and anticlockwise (decrease) rotational action.

tip trajectories of the two operators are both improved in the guidance mode. Fig. 8 also shows the operator's action detected by the master controller. The comparative results between Fig. 8 (b) and (d) with Fig. 8 (f) and (h) show that OA is reduced in the guidance mode. It indicates that the operation gains more fluency under the guidance of the planned path. Overall, the objective evaluation results demonstrate that the guidance of the planned path significantly improved the operator's operations.

For subjective assessment, the experimental protocol uses a seven-level Likert scale [38] to record metrics. The operators are asked to compare the guidance mode and free mode on a scale from -3 to +3 (-3: negative evaluation, +3: positive evaluation). The results are listed in Table 5. The majority of the users prefer the guidance mode over the free mode. The users feel more in control with guidance enabled, giving them a safer feeling. The visual guidance of the real-time planned path is evaluated as useful by the users. In addition, the *p* value of the significance test is smaller than 0.05. It indicates an overall better task performance with the guidance of the real-time planned surgical path.

5. Discussion

We present a surgical GAN framework for path planning of the passive flexible guidewire in ES. It directly takes the surgical state image and historical tool trajectory as input and outputs local tool path in real time. Both the unstructured surgical environment and unpredictable deformation of passive flexible tools rose prohibitive challenge for path planning by directly using explicit mathematical models. The proposed method solves this challenge by constructing a novel GAN architecture, of which the Generator and the Discriminator are both constructed by combining CNN and LSTM. In this way, the proposed surgical GAN has the capability to extract

and fuse the spatial features of medical images and temporal features of the historical tool trajectory. Fusion of the feature of medical images and historical tool trajectory can achieve a better representation of surgeons' operating experience for surgical path planning, which is indicated by Table 3. Although deep learning-based path planning method needs large scale of training sample, transfer learning method can be adopted in the future work to make use of the clinical samples and the samples collected with vascular models. In this way, the demanded number of clinical samples can be reduced.

This work is an important primary step to turn automatic real-time surgical path planning into clinical practice for ES. It has great potential to be translated to real-life conditions. Because deformable vascular models are costly, we made 10 vascular models for data collection and algorithm test by sculpting the vascular contours in 10 boards that cast in beeswax. It means that the vascular models are rigid and can not deform in the experiments. In the actual vascular system, the guidewire and the vascular contours can both deform. However, for one hand, the guidewire within the experimental stand can actually deform when interacting with the inner walls of the vascular models. For the other hand, the local images near the guidewire tip are used to represent the surgical state and input to the network. The local images are obtained in real time by cropping the region near the guidewire tip with a certain size. During the guidewire insertion, the scenes in the local images input to the network are changeable. It means that the shapes and relationship of the vascular contours and the deformable guidewire in the local images are dynamic and changeable. So, the proposed framework could be translated to real-life conditions with some extension. Firstly, for image preprocessing, we adopted the VRBC-t-TNN vascular edge detection method developed by Qin [32] to extract the contours of the vessels and guidewire. This algorithm has been demonstrated to be an effective method for accurately recovering vessel structures from the X-ray angiography images of moving organs or tissues, which is suitable for deformable vessels extraction in real-life conditions. Secondly, for the proposed surgical GAN, the inputs are preprocessed image of the contours of the vessels and guidewire. It means that the structure of the surgical GAN do not need change when translate to real-life conditions. Finally, it is true that inser-

Table 5
The Qualitative Results of User Experiment [-3, +3].

Metrics (mm)	Free mode	Guidance mode	<i>p</i> value
Performance efficiency	-2.05	1.55	0.02 < 0.05
Performance safety	-1.5	1.2	
Frustration	-1.7	1.55	
Overall performance	-1.45	1.28	

tions in deformable vessels are not contained in our training data, which should be compensated for translate the surgical GAN to real-life conditions. In the next research work, we will use deformable vessels like animal vessels or even human vessel in clinics, together with C-arm angiography machine, for further sample data collection. Transfer learning methods will be adopted to make use of the sample datasets collected in both deformable vessels and rigid vascular models.

Furthermore, a visual-haptic feedback mechanism can be further researched to provide enhanced guidance for the operators. We have previously investigated magnetorheological fluids (MR) based haptic interfaces for the master-follower ES robot [4]. Haptic feedback could be introduced and combined with visual feedback in the future work. Then a visual-haptic feedback for the operators could provide better guidance.

In addition, although only a guidewire is preliminarily used for path planning research in this work, our algorithm could be further extended to support simultaneous path planning of multiple tools. It is known that an ES task always needs collaborative operation of more than one tool (like the guidewire and the catheter) to achieve better flexibility. We consider that the proposed framework could be adapted for multiple tools' path planning, by adding historical tool paths and future paths of other tools respectively as the input and output of current network. In our future work, a further framework that transfer well across tools and across tasks would be studied to achieve broad applicability.

We would like to stress the benefit of the proposed methodology and its potential applications in surgeons' skills training and ES robotic autonomy. First, the objective results of user experiments demonstrate that the visual guidance of surgical path improves significantly users' operation from the point of the metrics of OD, TD, OA and TT. Although the planned path of the guidewire tip can not be accurately followed by the surgeon, due to the flexibility and unpredictable deformation of the guidewire body, the planned path can provide the surgeon with visual guidance, especially during the training of novice surgeon. Because the path planning is realized by learning a large amount of demonstrations of expert surgeons, their experience contained in the demonstration data is embedded in the planned path. Under the real-time visual guidance of the planned path, the trainee can adjust their surgical action in every step to obtain a better operation training. It is similar to the hand-holding training mode and beneficial for the training of novice surgeons. In the future work, an assessment model of the operation trajectory can be developed to provide improved impact to the trainees. Quantitative assessment could allow the trainees to better diagnose their skill deficits in real time compared with expert surgeons. It could enhance the learning efficiency, and reduce the long training period.

Second, automatic path planning lays an important base for improving the autonomy of current master-follower ES robots. In our future research, a visual servo control algorithm for supervised-autonomous robotic insertion will be developed, where the robot will drive the guidewire tip (through the guidewire end) to follow the next path point in every servo control cycle, according to the relationship between the position of the current guidewire tip and the position of the next path point. Supervised autonomous robotic suturing surgery has demonstrated its better efficacy, safety, and access to the best surgical techniques regardless of different human factors [2]. In our previous work [15], a trained convolutional neural network (CNN) is directly used for automatic navigation of basic guide wire insertion tasks in a Y-shaped vascular model. However, it is hard to assure the reliability of the action decision in this manner, due to the black-box problem and inexplicability of deep neural networks. The automatically planned path can provide a media for the surgeon to understand and supervise the autonomous robotic insertion.

6. Conclusions

In this paper, we propose a surgical GAN framework for path planning of passive flexible guidewire in ES, of which the generator is constructed by combining CNN and LSTM. It has the capability of unstructured surgical environment perception and passive flexible tool path planning in ES.

- (1) The results of comparison tests between the proposed surgical GAN, baseline networks and the non-rigid registration method indicate that the former can realize effective extraction and fusion of the spatial features in surgical state image and temporal features in historical tool trajectory. In this way, the tool path planning accuracy can be significantly improved.
- (2) The user experimental results demonstrate that the guidance of real-time planned surgical path significantly improves the operators' performance.
- (3) The results of this work motivate further studies including: enhancing the clinical practice of surgical GAN by clinical samples, quantitatively assessing the trainees' skill deficits and improving their training efficiency, and improving autonomy of ES robots.

CRedit authorship contribution statement

Yan Zhao: Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Yuxin Wang:** Methodology, Software, Writing - review & editing. **Jianhua Zhang:** Conceptualization, Methodology, Supervision. **Xinke Liu:** Data curation. **Youxiang Li:** Supervision. **Shuxiang Guo:** Conceptualization, Supervision. **Xu Yang:** Data curation. **Shunming Hong:** Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We acknowledge the efforts devoted by surgeons of Beijing Tiantan Hospital and 3rd Medical Center of Chinese PLA General Hospital for collecting training samples and providing an operation room for experiments.

Funding: This work was supported by the National Nature Science Foundation of China [No.62003128], Natural Science Foundation of Hebei Province of China [No.F2020202053], Natural Science Foundation of Tianjin City of China [No.20 J

CYBJC00610] and Preferential funding for postdoctoral research projects of Hebei Province of China [No.B2020003020].

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.neucom.2022.05.044>.

References

- [1] Louaï Adhami, Ève Coste-Manière, Optimal planning for minimally invasive surgical robots, *IEEE Trans. Robot. Autom.* 19 (5) (2003) 854–863.
- [2] A. Shademan, RS Decker, JD. Opfermann, et al. Supervised autonomous robotic soft tissue surgery. *Sci Transl Med*, 8(337):337ra64, 2016..
- [3] L. Napalkovaa, J.W. Rozenblit, G. Hwang, et al., An optimal motion planning method for computer-assisted surgical training, *Appl. Soft Comput.* 24 (2014) 889–899.

- [4] S. Guo, Y. Song, X. Yin, et al., A novel robot-assisted endovascular catheterization system with haptic force feedback, *IEEE Trans. Robot.* 35 (3) (2019) 685–696.
- [5] Y. Zhao, S. Guo, N. Xiao, et al., Operating force information on-line acquisition of a novel slave manipulator for vascular interventional surgery, *Biomed. Microdevices* 20 (2018) 33.
- [6] D. Huang, P. Tang, Y. Wang, et al., Computer-assisted path planning for minimally invasive vascular surgery, *Chin. J. Electron.* 27 (6) (2018) 1241–1249.
- [7] Y. Cheng, A. Firouzmanesh, A. Leleve, et al. Enhanced segmentation and skeletonization for endovascular surgical planning. In *Proceedings of SPIE*, page 83162W, 2012..
- [8] Y. Zhao, S. Guo, Y. Wang, et al., Vessel enhancing for a continuous dsa method towards endovascular interventional surgery, in: *IEEE International Conference on Mechatronics and Automation*, 2018, pp. 608–613.
- [9] G. Fagogenis, M. Mencattelli, Z. Machaidze, et al., Autonomous robotic intracardiac catheter navigation using haptic vision. *Science, Robotics* 4 (29) (1977) 2019.
- [10] H. Cao, M.A. Speidel, J. Tsai, et al., Fem analysis of predicting electro-myocardium contact from rf cardiac catheter ablation system impedance, *IEEE Trans. Biomed. Eng.* 49 (6) (2002) 520–526.
- [11] G. Wang, W. Li, M.A. Zuluaga, et al., Interactive medical image segmentation using deep learning with image-specific fine tuning, *IEEE Trans. Med. Imaging* 37 (7) (2018) 1562–1573.
- [12] H. Gupta, K.H. Jin, H.Q. Nguyen, et al., Cnn-based projected gradient descent for consistent ct image reconstruction, *IEEE Trans. Med. Imaging* 37 (6) (2018) 1440–1453.
- [13] T. Kim, J. Heo, D.K. Jang, et al..
- [14] Y. Zhao, S. Guo, Y. Wang, et al., A cnns-based prototype method of unstructured surgical state perception and navigation for an endovascular surgery robot, *Med. Biol. Eng. Comput.* 57 (2019) 1875–1887.
- [15] P. Li, Z. Yang, S. Jiang, Needle-tissue interactive mechanism and steering control in image-guided robot-assisted minimally invasive surgery: a review, *Med. Biol. Eng. Comput.* 56 (6) (2018) 931–949.
- [16] R. Alterovitz, M. Branicky, K. Goldberg, Motion planning under uncertainty for image-guided medical needle steering, *Int. J. Robotics Res.* 27 (11–12) (2008) 1361–1374.
- [17] E. Dehghan, S.E. Salcudean, Needle insertion parameter optimization for brachytherapy, *IEEE Trans. Robot.* 25 (2) (2009) 303–315.
- [18] V. Duindam, J. Xu, R. Alterovitz, et al., 3d motion planning algorithms for steerable needles using inverse kinematics, *Int. J. Robotics Res.* 57 (2009) 535–549.
- [19] S.P. DiMaio, S.E. Salcudean, Needle steering and motion planning in soft tissues, *IEEE Trans. Biomed. Eng.* 52 (6) (2005) 965–974.
- [20] K.B. Reed, A. Majewicz, V. Kallem, et al., Robot-assisted needle steering, *IEEE Robot. Autom. Mag.* 18 (4) (2011) 35–46.
- [21] J. Fauser, G. Sakas, A. Mukhopadhyay, Planning nonlinear access paths for temporal bone surgery, *Int. J. Comput. Assist. Radiol. Surg.* 13 (5) (2018) 637–646.
- [22] D. Baek, M. Hwang, H. Kim, et al., Path planning for automation of surgery robot based on probabilistic roadmap and reinforcement learning, in: *International Conference on Ubiquitous Robots*, 2018, pp. 342–347.
- [23] J. Schulman, J. Ho, C. LEE, et al. *Learning from Demonstrations Through the Use of Non-Rigid Registration*, page 339–354. Springer, 2016..
- [24] L. Geert, K. Thijs, E.B. Babak, et al., A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [25] N. Nasr-Esfahani, E. and Karimi, Jafari M.H., et al. Segmentation of vessels in angiograms using convolutional neural networks. *Biomed. Signal Process. Control*, 40:240–251, 2018..
- [26] D. Nie, R. Trullo, J. Lian, et al., Medical image synthesis with deep convolutional adversarial networks, *IEEE Trans. Biomed. Eng.* 65 (12) (2018) 2720–2730.
- [27] Y. Zeng, X. Liu, N. Xiao, et al., Automatic diagnosis based on spatial information fusion feature for intracranial aneurysm, *IEEE Trans. Med. Imaging* 39 (5) (2020) 1448–1458.
- [28] T. Fernando, S. Denman, S. Sridharan, et al. Learning temporal strategic relationships using generative adversarial imitation learning. *arXiv preprint arXiv:1805.04969*, 2018..
- [29] T. Fernando, S. Denman, S. Sridharan, et al., Soft+hardwired attention: An lstm framework for human trajectory prediction and abnormal event detection, *Neural Netw.* 108 (2018) 466–478.
- [30] A. Gupta, J. Johnson, F. Li, et al. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Conference on Computer Vision and Pattern Recognition*, number 1, pages 1–10, 2018 2018..
- [31] V. Oriol, T. Alexander, B. Samy, et al. Show and tell: A neural image caption generator. *CoRR*, abs/1411.4555, 2014..
- [32] B. Qin, M. Jin, D. Hao, et al., Accurate vessel extraction via tensor completion of background layer in x-ray coronary angiograms, *Pattern Recognit.* 87 (2019) 38–54.
- [33] T.Y. Zhang, C.Y. Suen, A fast parallel algorithm for thinning digital patterns, *Commun. ACM* 27 (3) (1984) 236–239.
- [34] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., Generative adversarial nets, in: *Conference and Workshop on Neural Information Processing Systems*, 2014.
- [35] K. Clark, M.T. Luong, V.Q. Le, et al. Electra: Pre-training text encoders as discriminators rather than generators. *ICLR*, 2020..

- [36] X. Bao, S. Guo, N. Xiao, et al., Operation evaluation in-human of a novel remote-controlled vascular interventional robot, *Biomed. Microdevices* 20 (2018) 34.
- [37] S. Oonsiri, C. Jumpangern, T. Sanghangthum, Radiation dose to medical staff in interventional radiology, *J. Med. Assoc. Thai* 90 (4) (2007) 823–828.
- [38] V.R. Preedy and R.R. Watson, editors. *Likert Scale*, pages 4248–4248. Springer, New York, 2010..



Yan Zhao was born in Hebei, China, in 1988. He received the B.E. and the M.E. degrees in mechanical engineering from Hebei University of Technology, China, in 2012 and 2015, respectively, and the Ph.D. degree in biomedical engineering from Beijing Institute of Technology, China, in 2019. He is currently a post-doc in Innovation Team in Key Areas of the Ministry of Science and Technology, Hebei University of Technology. His current research interests include surgical robotics, medical image processing, deep learning, robot learning.



Yuxin Wang was born in Anhui, China in 1996. He is currently pursuing the master's degree in biomedical engineering with School of Life Sciences, Beijing Institute of Technology, China. He received the B.Sc in automation from Beijing Institute of Technology, China. His research interests include surgical robotics, machine learning and biomedical engineering.



Jianhua Zhang was born in Hebei, China, in 1979. He received the B.E. degrees in Mechanical Engineering from Agricultural University of Hebei, China, in 2003, and the M.E. and Ph.D. degrees in Mechanical Engineering from Shandong University, China, in 2005 and 2008 respectively. He is currently a professor and the leader of Innovation Team in Key Areas of the Ministry of Science and Technology, Hebei University of Technology. His research interests include surgical robots, rehabilitation exoskeleton.



Xinke Liu was born in 1989. He received the B.Sc degree in Hebei Medical University in 2014 and M.sc in Capital Medical University in 2014. He received his resident training in Beijing Tiantan Hospital, Capital Medical University. He is currently a resident doctor at Beijing Tiantan Hospital. His major interest including endovascular surgery robotic system clinical experiment and black blood imaging and hemodynamic study of intracranial aneurysms.



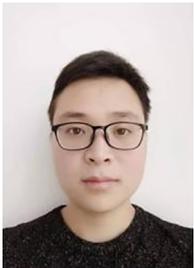
Youxiang Li was born in 1966. In 1990, he received his B.Sc in Tongji Medical University; in 1998, he received M.sc in Capital medical university. He received his resident and fellow training in Beijing Tiantan Hospital, Capital Medical University. In 2009, he became the chief of Department of Interventional Neuroradiology, Beijing Neurosurgical Institute and Beijing Tiantan Hospital. His research interests include endovascular surgical robot, vessel wall imaging and hemodynamic study of cerebral vascular disease.



Shuxiang Guo was born in Jilin, China in 1963. He received the Ph.D. degree from Nagoya University in 1995. He is currently a professor in Biomedical engineering at Beijing Institute of Technology and Kagawa University, Japan. He serves as an academican of the Japanese Academy of Engineering and an IEEE Fellow. He is the leader of Key Laboratory of Convergence Biomedical Engineering System and Healthcare Technology, the Ministry of Industry and Information Technology, China. He is the Editor in Chief of *IJMA*. His research interests include surgical robotics, rehabilitation robotics and endoscopic capsule robotics.



Shunming Hong was born in Chongqing, China, in 1989. He is a neurosurgeon and engaged in neurovascular intervention in the Third medical centre, Chinese PLA(People's Liberation Army) General Hospital in Beijing, China. He received the master of medicine from Tianjin medical University, China, in 2016. His research interests include interventional treatment of cerebrovascular and spinal vascular diseases, interventional surgical robotics and biomedical engineering.



Xu Yang was born in Henan, China in 1997. He is currently studying for a master's degree in mechanical engineering at the School of Mechanical Engineering, Hebei University of Technology, China. He received a bachelor's degree in mechanical design and manufacturing and automation from Henan University of Technology in China. His research interests include surgical robotics and machine learning.