Proceedings of 2021 IEEE
International Conference on Mechatronics and Automation
August 8 - 11, Takamatsu, Japan

Study on the Automatic Surgical Method of the Vascular Interventional Surgical Robot Based on Deep Learning

Jian Guo[1], Suxiang Feng[1]

[1]Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems and Intelligent Robot Laboratory
Tianjin University of Technology
Binshui Xidao 391, Tianjin, China

jianguo@tjut.edu.cn; 1549158571@qq.com;

Shuxiang Guo[1,2*]

[2] Department of Intelligent Mechanical Systems Engineering, Faculty of Engineering,
Kagawa University
2217-20, Hayashi-cho, Takamatsu 761-0396, Japan

*Corresponding author: guo@eng.kagawa-u.ac.jp;

***Abstract* - Because of the various shortcomings of traditional interventional surgery, it has become a trend to develop vascular interventional robots to assist doctors in completing operations. However, the existing vascular intervention robot assist system has given the doctor a large amount of data to assist the operation, but the doctor still needs a lot of energy and physical strength to operate the handle at the master side. Therefore, this paper aims to study the use of deep learning algorithms in a laboratory environment to assist the vascular interventional surgery robot to drive the catheter to the specified position. In the future, using this method can effectively reduce the burden on doctors and reduce operating errors caused by doctors' fatigue.**

Index Terms -Vascular interventional surgical robot system, Deep learning, Auxiliary system, Automatic surgery.

I. INTRODUCTION

With the acceleration of social aging and urbanization and the prevalence of unhealthy lifestyles of residents, the number of cardiovascular disease patients in the country has increased significantly. It is estimated that the number of currents and potential patients in China is about 290 million[1]. The most direct and effective method for vascular diseases is minimally invasive vascular interventional surgery. Minimally invasive vascular intervention surgery refers to the operation of the catheter/guidewire to the lesion by pushing, pulling, and twisting the catheter under the guidance of medical images to complete the operation for related diseases. In traditional vascular interventional surgery, doctors need to use an X-ray imaging system to complete the intubation operation in the operating room. Long-term accumulation of radiation will cause harm to the doctor's body; the interventional surgery method is highly skilled, which makes the interventional surgeon's training time long and high concentration[2]. The operating doctor is prone to fatigue and affects the quality of the operation. Due to the low accuracy of manual intubation, the combination of robotic technology and minimally invasive surgery has huge market demand for the application of interventional surgery robots in clinical practice[3].

In 2016, Corindus Vascular Robotics of the United States developed a new vascular interventional robot system named CorPath GRX on the original basis[4]. The system has an extension arm and a touch screen display, and is equipped with an automatic catheter guidance function, which optimizes the workflow greatly improves the clinical capabilities of the system. In 2018, the team of Professor Shuxiang Guo from the Beijing Institute of Technology developed a clinically applicable vascular interventional surgical robot system. The main side of the system uses two Touch X devices from Geomagic to control the catheter and the guidewire separately, and provide real-time force. Feedback function, and real-time monitoring of the operation situation through visual assistance, with good safety and stability[5]. In 2016, researchers from the Children's National Health System in Washington, USA developed the robot STAR (Smart Tissue Autonomous Robot). The research team performed two surgical operations with the STAR robot, connecting two sections of pig intestines together, one outside the body and one inside the living body. Both operations were successful. The pigs used for the operation did not have any complications[6].

However, doctors who concentrate on operating the vascular interventional surgical robot for a long time are also prone to fatigue and misjudgment. Therefore, it is necessary to establish an intelligent judgment and manipulation mechanism combined with the rich experience of the doctor to help the doctor reduce the workload in the operation and reduce the fatigue[8].

This paper mainly uses deep learning in a laboratory environment to realize the pre-judgment and automatic operation of the mechanical operation from the end, and analyzes the error through experiments to determine the feasibility. This paper is organized as follows: The first part introduces the research situation of vascular interventional surgery robot; the second part shows our team's vascular intervention robot platform; the third part describes the automatic operation method of the vascular interventional robot; the fourth part is to judge the feasibility of the method through experiments. And analyze the error; the fifth part, summarize the full text.

II. THE OVERVIEW OF PLATFORM

The vascular interventional operation platform of our team is divided into a master side and a slave side, where the master includes a master manipulator, a master controller, and a display screen; the slave side includes a slave manipulator, a slave controller, and a camera. The entire system operation process: the manipulator controls the master manipulator, and the master manipulator collects the manipulator's axial displacement and radial rotation information[9]. The collected information is processed by the master controller and the

978-1-6654-4098-1/21/$31.00 ©2021 IEEE

information is input to the slave controller. The slave controller controls the slave manipulator to control the catheter and the guidewire according to the received information. During the operation, the slave manipulator can collect the resistance information about the catheter and the guidewire and send it to the slave controller of processing, and then feed it back to the master side. The master controller controls the force feedback damper on the master manipulator to provide tactile feedback on the doctor based on the received feedback information[10]. During the operation, the slave side is also equipped with a camera to monitor the scene in real-time. The manipulator can observe the actual operation of the catheter and guidewire through the display screen on the master side, providing visual feedback to the doctor[11]. During the operation, the doctor observes the actual movement of the guidewire and catheter-based on the DSA(Digital subtraction angiography) image and judges the subsequent operation steps，the process is shown in Fig. 1.

However, doctors who concentrate on operating the vascular interventional surgery robot for a long time are also prone to fatigue[12].Therefore, it is necessary to establish an operation method judgment mechanism independent of the doctor to assist the doctor in the judgment to complete the operation. Therefore, based on the platform built by the team, our team proposed an automatic operating system for vascular interventional surgery based on deep learning to reduce the workload of doctors and achieve the purpose of assisting doctors.

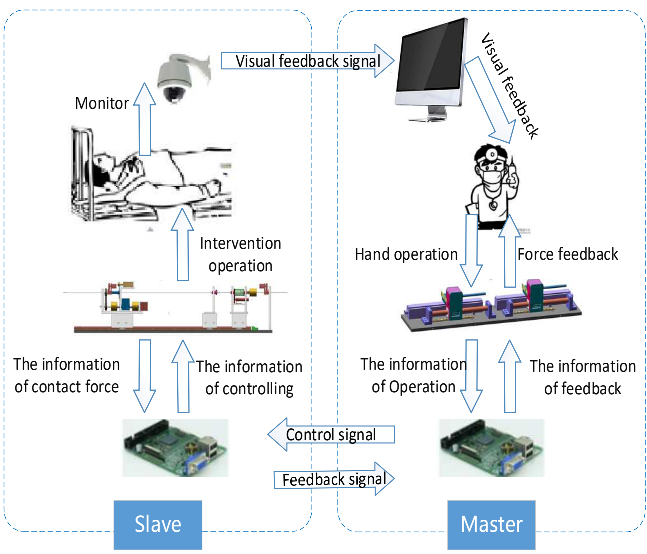

Fig. 1 Concept picture of vascular interventional surgical robot

III. The Automatic Surgery Plan for Vascular Interventional Surgical Robot

In order to practically develop an interventional robot automatic operating system that can assist doctors in completing vascular surgery, the main research goals of this paper are as follows: 1. Establish a judgment mechanism for operating techniques independent of doctors. Combine the judgment mechanism with the slave controller Complete automatic operation.

In order to accomplish the above purpose, this paper adopts a supervised learning method. First, we must establish a training set. The input items in the training set are medical images and the output items are the control signals of the controller. In the acquisition phase of the training set, the corresponding relationship between the medical image and the control signal is established, and the complicated continuous surgical operation and judgment problem is simplified to the classification problem of the medical image and the control signal. Compared with other classification problems, it has the following characteristics: 1. It does not pay attention to the color of the image and can be satisfied by using grayscale images; 2. The real-time requirements are higher, so higher execution efficiency and lower calculation amount are required. Aiming at the above-mentioned problems and characteristics, a vascular intervention surgery robot automatic surgery program based on the Alexnet deeping learning network was developed. Before training the Alexnet deeping learning network, we first need to preprocess the data in the training set for unified specifications, grayscale, edge extraction. Then use the preprocessed data set to train the Alexnet deeping learning network and export the trained model, and finally, use the trained model to judge the operation technique and output the control signal to the slave side controller to control the slave side mechanism to drive the guide wire catheter to move. The A schematic diagram of Automatic surgery plan for vascular interventional surgical robot is shown in Fig. 2.

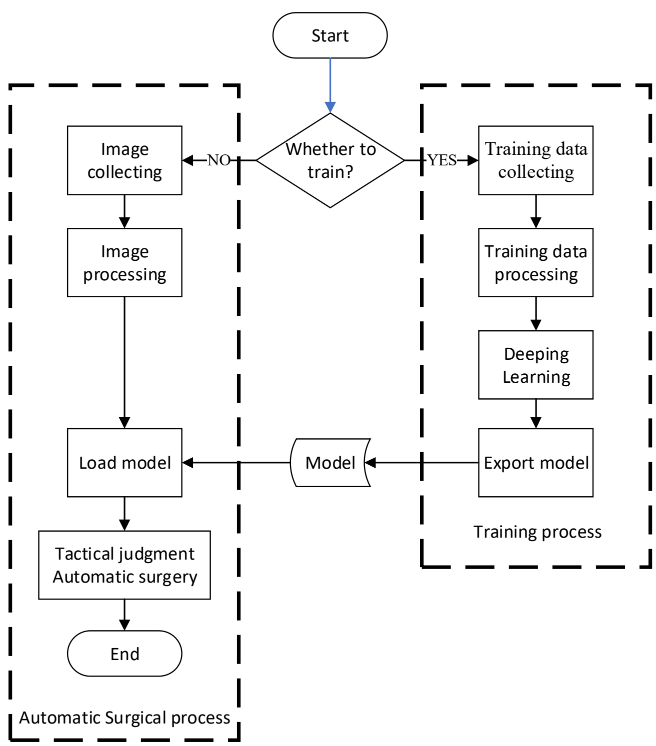

Fig. 2 The A schematic diagram of Automatic surgery plan for vascular interventional surgical robot

A. Collecting and processing for training data

In order to train the deep learning network model, a large number of images containing the posture of the catheter guidewire in the blood vessel are required as a training set. The size of the training set determines the accuracy of the model. The data set used in this paper is mainly processed by sampling continuous images and continuous corresponding control signals. Since this process can be regarded as a Markov process approximately, so we random it during the process of storing it in the data set[13].

There are two main types of training set processing. One is image graying and edge extraction, and the second is to number control signals. In the RGB image model, the image is usually superimposed by three-color images[14]. In the paper, we only need to use the gray image to meet the training requirements. The process is shown in formula (1):

$$Gray = \sqrt[2.2]{\frac{R^{2.2} + (1.5G)^{2.2} + (0.6B)^{2.2}}{1 + 1.5^{2.2} + 0.6^{2.2}}} \qquad (1)$$

Then we can further perform edge extraction based on the grayscale image. This paper uses the canny edge detection algorithm. The Canny algorithm is mainly divided into three steps: the first step is to filter the main purpose of Gaussian filtering is to reduce noise, general image processing algorithms need to perform noise reduction first, and filtering mainly makes the image smooth. The Gaussian function is a function that is similar to the normal distribution with the larger and smaller sides in the middle. For a pixel at a position *(m, n)*, its gray value is *f (m, n)*[14], then the gray value after Gaussian filtering will be as shown in the formula (2):

$$G_\sigma(m,n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{m^2+n^2}{2\sigma^2}} . f(m,n) \qquad (2)$$

After the image denoising algorithm, the Sobel operator is used to detect the horizontal gradient *G (m, n)* and vertical gradient θ of the denoised image, and then the formula (3) is used to calculate the image gradient modulus length *G* and the gradient direction θ[14].

$$\begin{gathered} G(m,n) = \sqrt{g_x(m,n)^2 + g_y(m,n)^2} \\ \theta = arctan\frac{g_x(m,n)}{g_y(m,n)} \end{gathered} \qquad (3)$$

But in the process of Gaussian filtering, the edges may be magnified. Dual-threshold non-maximum suppression is needed to filter the points that are not edges, so that the width of the edge is as corresponding as possible. The specific method is to set two thresholds, a high threshold, and a low threshold. If the gradient value of a pixel is greater than the high threshold, the pixel is directly determined as an edge. When the gradient value of a point is between the high threshold and the low threshold, the pixel is determined as a pending edge point. If the gradient of a point is less than the low threshold, the point will be suppressed, as shown in the formula (4):

$$M_T(m,n) = \begin{cases} M(m,n) & M(m,n) > T \\ M(m,n)' & t \le M(m,n) \le T \\ 0 & M(m,n) < T \end{cases} \qquad (4)$$

The picture after the image processing is shown in Fig. 3.

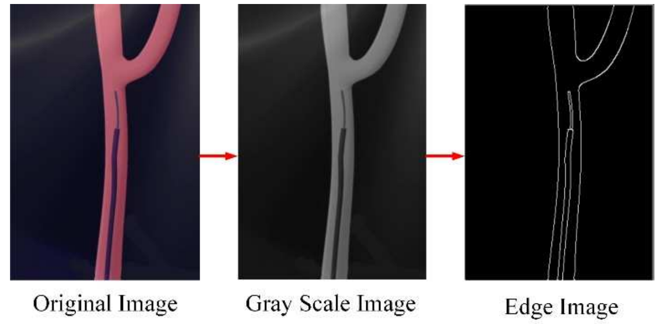

Fig. 3 Image processing diagram

The main side operation information is mainly divided into five types: forward, backward, left, right, and standstill signals. We extract the enable signals of the four signals on the slave controller while extracting the image data. The control signals correspond to the images at the same time one by one, and finally form a complete training data.

B. Deep Learning process

Alexnet is one of the classic deep learning network models，it has excellent nonlinear fitting capabilities and image processing capabilities. It consists of input layer, convolutional layer, pooling layer, activation function layer, and fully connected layer. The structure of Alexnet is shown in Fig. 4.

Each neuron in alexnet only needs to perform convolution operation with the data in the corresponding local block in the image to obtain the characteristics. This local connection method can greatly reduce the number of parameters involved in the calculation in the network and increase the calculation speed . There may be similar local features between different images, and similar features may also exist in different regions of the same image. Then using the same convolution kernel to perform convolution calculations on the entire image will not affect the final features. Therefore, the weight parameters can be shared to further reduce the network parameters[15].

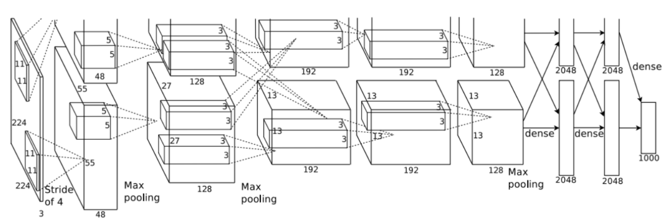

Fig. 4 The structure of Alexnet

The feature extraction part of the core is completed by the convolutional layer. The main work of the convolutional layer is to perform convolution operations with the input image through the convolution kernel to obtain the description of the feature, that is, the feature map. Multiple convolution kernels can obtain corresponding multiple feature maps through calculation. These convolution kernels can automatically learn the pixel-level features of the image, so the convolution layer saves a lot of human resources[16]. The calculation formula of convolution operation is shown in formula (5):

$$x_j^l = f\Big(\sum_{j\in M^{l-1}} x_i^{(l-1)} * k_{ij}^l + b_j^l\Big) \qquad (5)$$

x_j^l represents the feature *map j* of l*ayer l*, k_{ij}^l represents the convolution kernel connecting the feature *map j* of the *layer l* and the feature *map i* of the *layer l-1* , M^{l-1} represents the input feature atlas selected by the *layer l-1*, b_j^l represents the bias term, $*$ represents convolution operation, *f()* represents a nonlinear activation function

The principle of pooling is to divide the feature map obtained in the convolutional layer into blocks in a certain way, and then use the calculated value on each block to represent the characteristics of the entire region, which is a special convolution operation. Therefore, the pooling operation can further reduce the number of network parameters and reduce the feature dimension[16]. The steps of convolution and pooling are shown in Fig. 5

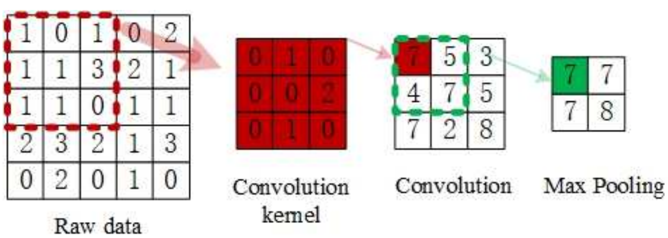

Fig. 5 Schematic diagram of convolution and pooling

Finally, the fully connected layer is generally located at the top of the network to identify the abstract features learned by the convolutional layer and output the classification results. Its neurons use all connections, and the number is the same as the number of operations[17]，as shown in the formula(6):

$$y_{ij}^l = softmax\Big(\sum_{j\in M^{l-1}} x_i^{(l-1)} * w_{ij}^l + b_j^l\Big) \qquad (6)$$

softmax() is the activation function, y_{ij}^l is the output after training, w_{ij}^l is weight coefficient, $x_i^{(l-1)}$ is feature input, b_j^l represents the bias term, *i, j* represents index of the parameter in two-dimensional coordinates.

Finally, using the cross-entropy loss function to measure the degree of model fit[17], the calculation formula is formula (7):

$$C = -\frac{1}{m}\sum_{i=1}^{m} y^i \log(y^i) + (1-y^i)\log(1-y^i) \qquad (7)$$

m is the number of categories to be classified, y^i is the true expected label of sample i. C is the output label obtained through network training.

C. *Automatic surgical process*

In the process of automatic surgery, we first collect real-time images of blood vessels and the end of the guide wire catheter during the operation and import the images into the completed training model after preprocessing to obtain the operation mode judged by the deep learning network. The upper computer will the control signal corresponding to the operation mode is transmitted to the slave controller, and the slave controller drives the slave mechanism to deliver the guidewire catheter to the designated position. Repeat the above until the catheter reaches the designated position. The control flow is shown in Fig. 2

IV. Experiments And Results

A. *The process of experiment*

Before verifying automatic surgery, we first need to collect the training set data required by the Alexnet deep learning according to the schematic as shown in Fig. 2.

Our operators operate on the main side operator according to the medical images taken by the camera. The upper computer simultaneously captures the image data and the corresponding operation signals, and the sampling interval of each segment is set to 0.25 seconds, as shown in Fig. 6. In order to prevent the occurrence of over-fitting, we also built 9 groups of different blood vessel models to operate to obtain different operating data, as shown in Fig. 7.

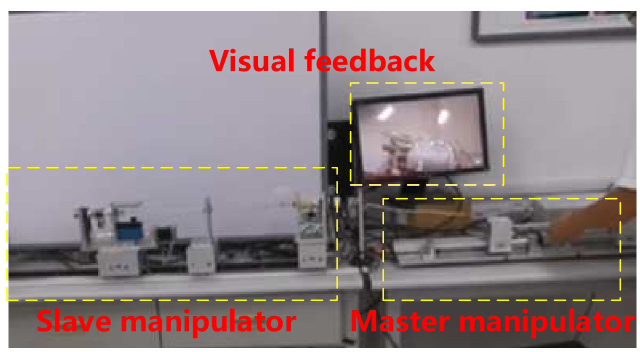

Fig. 6 The sampling process

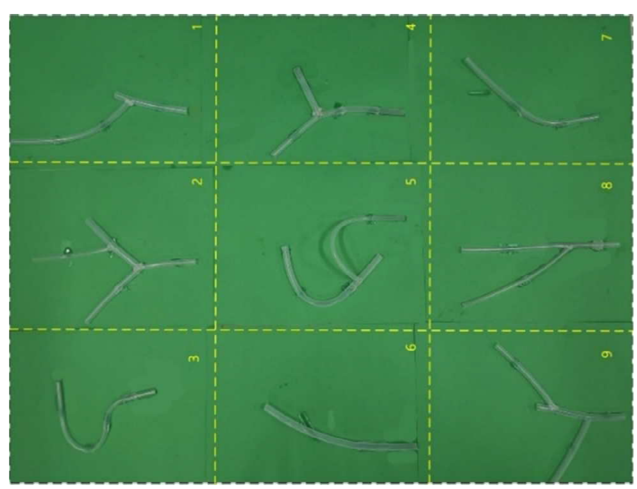

Fig. 7 The blood vessel model

After obtaining the image data and operation data, we preprocess them. The image data is processed by two methods of grayscale and canny edge calculation, which are respectively reserved as grayscale image and edge algorithm processing image; the control signal forward code is 1000, backward code is 0100, left turn code is 0010, right turn code is 0001 and the static code is 1111.

Input the training data into the alexnet deep neural network for training. The operating hardware platform of Alexnet is cpu intel i5-9300h 2.4ghz, gpu grx1650; the operating environment of the software level upper computer control part is Windows 10 64bit, the integrated development environment is VistualStudio Code, the programming language is Python, and the building Alexnet network in Tensorflow environment Train the network according to the training process in the block diagram, and output the network model after the training is completed.

During the operation process, the medical image of the operation process is first captured by the medical camera, and the medical image is captured through the win32 screen capture program every 0.25 seconds, and the size of the medical image is compressed, and then the Opencv library is called to enhance the edge features of blood vessels and catheters extract. The preprocessed medical images are imported into the trained deep learning network, and the deep learning network will output predicted operation codes according to the trained network rules. The upper computer compiles the operation code into a control signal and transmits it to the Arduino microprocessor. The Arduino microprocessor then converts the control signal into a PWM wave control signal to control the ARM24SAK-H100 stepping motor to perform a specified movement within 0.25 seconds, stepping the motors to drive the slave side to push the catheter to continuously complete the designated action until the catheter reaches the predetermined position for the operation. We use NDI to tie the catheter to track the position of the catheter end in real time, as shown in the Fig. 8.

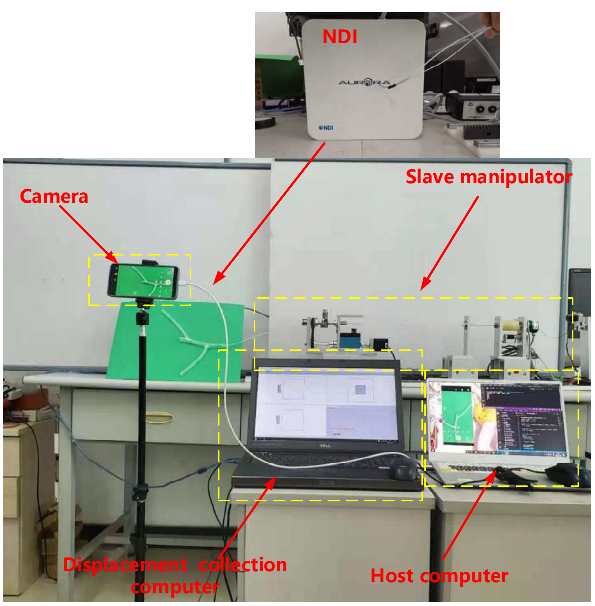

Fig. 8 The automated surgery and data collection process

B Example1

When walking in path 7, because the path is single, the average distance between ten paths is 1.538mm, and the ten paths of the end of the catheter is shown in the Fig. 9.

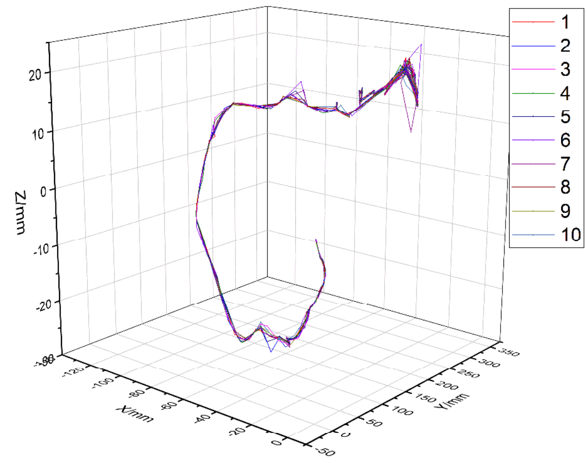

Fig. 9 the end paths of Catheter in path 7

C Example 2

When walking in path 2, because the path has a bifurcation, in the ten paths, two times go to bifurcation 1, one time to bifurcation 2, and the other 7 times go to the target point. The difference between the paths is shown in Fig. 10

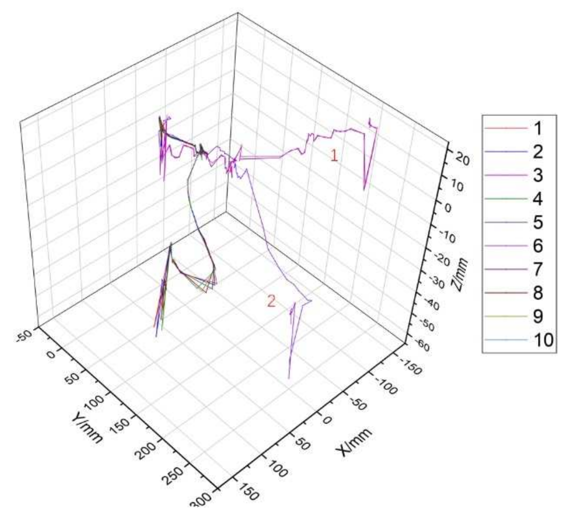

Fig. 10 the end paths of Catheter in path 2

In the multi-component fork path1, the success rate of reaching the target point is 80%, in the path2, the success rate of reaching the target point is 70%; in the path4, the success rate of reaching the target point is 80%; in the path5, the success rate of reaching the target point is 90%; in the path8, the success rate of reaching the target point is 70%; in the path9, the success rate of reaching the target point is 60%, The average success rate of multipath testing is 75%.

The conclusion proves that in the laboratory environment, the vascular access surgical robot can perform automatic surgery, but due to various constraints of the laboratory such as insufficient training set data, insufficient blood vessel types, low information sampling frequency, and other limiting factors, which result in the success rate cannot reach more than 90%. In the follow-up work, these constraints will be optimized one by one.

V. Conclusion And Future Work

In order to solve the problem of fatigue caused by doctors in the operation of vascular interventional robots, this paper proposed to use the deep learning network combined with images to make operation judgments, and according to the judgment made by the deep learning network, the automatic operation of the vascular intervention robot is realized. In the laboratory environment, after experimental analysis, this method can realize the automatic operation of a relatively simple path, but there is a certain error rate in the judgment of the complex path.

In the future research work, relevant steps will be optimized to reduce the occurrence of errors, and a judgment mechanism will be introduced to remind doctors of manual intervention to prevent accidents.

Acknowledgments

This research is supported by National Natural Science Foundation of China (61703305), Key Research Program of the Natural Science Foundation of Tianjin (18JCZDJC38500) and Innovative Cooperation Project of Tianjin Scientific and Technological (18PTZWHZ00090).

References

[1] Z.Wang, S.Hu , "Chinese Cardiovascular Disease Report 2019", *Chinese Journal of Hypertension*, vol. 25, no. 5 pp. 401-420, 2020.

[2] T Hu, T.Tian, K.Wang , L.Li, "Comparative study on the effect of craniotomy and intravascular interventional aneurysm embolization on the rupture of high-risk intracranial and eurysms in young and middle-aged people", *Practical Journal of Cardio-Cerebro-Pulmonary Vascular Disease*, vol. 23, no. 7, pp. 81-84, 2015.

[3] X.Qian, "Observation on the effect of endovascular interventional therapy in patients with failed hemostasis under endoscope", Chinese Health Standard Management, vol. 9, no. 11, pp. 52-54, 2018.

[4] Smitson C. C., Ang L., Pourdjabbar A., "Safety and Feasibility of a Novel, Second-Generation Robotic Assisted System for Percutaneous Coronary Intervention: First in Human Report", *Journal of Invasive Cardiology*, vol. 30, no.4, pp. 152-156, 2018.

[5] Y.Song, S.Guo, X.Yin, "Performance evaluation of a robot-assisted catheter operating system with haptic feedback", *Biomedical Microdevices*, vol. 20, no. 2, 2018.

[6] X.Bao, S.Guo, N.Xiao, " A cooperation of catheters and guidewires-based novel remote-controlled vascular interventional robot", *Biomed Microdevices*, vol. 20, no.1, pp.20, 2018.

[7] M.Deng, C.Lin, Z.Qin, C.Huang, "The value of three-dimensional digital subtraction angiography in the diagnosis of cerebrovascular diseases", *China Digital Medicine*, vol.12, no. 12, pp. 41-43, 2017.

[8] C.Meng, J.Guo,S. Guo,Q.Zhan, L.Qi, "Design and Evaluation of a Novel Slave Manipulator for the Vascular Interventional Robotic System" , *Proceedings of 2019 IEEE International Conference on Mechatronics and Automation*, pp.1350-1355, August 4.7,Tianjin, China, 2019.

[9] S.Guo, Q.Zhan, J.Guo, C.Meng, L.Qi, "A Novel Vascular Interventional Surgeon Training System with Cooperation between Catheter and Guidewire" ,*Proceedings of 2019 IEEE International Conference on Mechatronics and Automation*, pp. 1403-1408, August 4-7, Tianjin, China, 2019.

[10]J.Guo, X.Jin, S. Guo, "Study of the Operational Safety of a Vascular Interventional Surgical Robotic System", *Micromachines,* 9(3):119, 2018.

[11]J.Guo, Y.Cheng, S.Guo, W.Du, "A novel path planning algorithm for the vascular interventionalsurgical robotic doctor training system", *Proceedings of 2017 IEEE International Conference on Mechatronics and Automation*, pp. 45-50, 2017.

[12]K.Chen, "Research on key technologies of X-ray image guidance in vascular interventional surgery", University of Chinese Academy of Sciences (Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences), 2019.

[13]L.Chen."Research on Key Technologies of Image Processing in Assisted Vision ", Xi'an University of Electronic Technology, 2016.

[14]H.Qian, "Medical image edge detection algorithm based on improved Canny operator", *Software Guide*,vol.18(2), pp. 51-54, 2019.

[15]Krizhevsky A, Sutskever I, Hinton G, ImageNet classification with deep convolutional neural networks", *Neural Information Processing Systems*. pp.1097-1105, 2012.

[16]K He, X Zhang, S Ren, "Deep residual learning for image recognitiong ", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp* .770-778, 2016.

[17]N Zeng, H Zhang, "Facial expression recognition via learning deep sparse autoencoders", *Neurocomputing*, vol. 273: 643-649, 2018