

Document downloaded from:

<http://hdl.handle.net/10251/188311>

This paper must be cited as:

Huang, C.; Zong, Y.; Chen, J.; Liu, W.; Lloret, J.; Mukherjee, M. (2021). A Deep Segmentation Network of Stent Struts Based on IoT for Interventional Cardiovascular Diagnosis. *IEEE Wireless Communications*. 28(3):36-43.  
<https://doi.org/10.1109/MWC.001.2000407>



The final publication is available at

<https://doi.org/10.1109/MWC.001.2000407>

Copyright Institute of Electrical and Electronics Engineers

Additional Information

# A Deep Segmentation Network of Stent Structs based on IoT for Interventional Cardiovascular Diagnosis

Chenxi Huang, Yongshuo Zong, Jinling Chen, Weipeng Liu, Jaime Lloret, and Mithun Mukherjee

## ABSTRACT

The Internet of Things (IoT) technology has been widely introduced to the existing medical system. An eHealth system based on IoT devices has gained widespread popularity. In this article, we propose an IoT eHealth framework to provide an autonomous solution for patients with interventional cardiovascular diseases. In the framework, the wearable sensors are used to collect a patient's health data, which is daily monitored by a remote doctor. When the monitoring data is abnormal, the remote doctor will ask for image acquisition of the patient's cardiovascular internal conditions. We leverage edge computing to classify these training images by the local base classifier, thereafter, the pseudo-labels are generated according to its output. Moreover, a deep segmentation network is leveraged for the segmentation of stent struts in the intravascular optical coherence tomography (IVOCT) and intravenous ultrasound (IVUS) images of patients. The experimental results demonstrate that the remote and local doctors perform real-time visual communication to complete the telesurgery. In the experiments, we adopt the U-net backbone with a pretrained SeResNet34 as the encoder to segment the stent struts. Meanwhile, a series of comparative experiments have been conducted to demonstrate the effectiveness of our method based on accuracy, sensitivity, jaccard, and dice.

## INTRODUCTION

The Internet of Things (IoT) is a system generally comprised of a large number of distributed mobile terminals embedded with sensors. IoT applications typically range from several fields, such as sports, agriculture, commerce, households, and medical care. In recent years, IoT has developed rapidly as it can minimize the effort with high accuracy, and effectively reduce time by predicting the outlook in a better way. Integrating IoT devices with actuators and sensors into gadget-like appliances and machines can introduce automation to almost every field.

Healthcare is one of the most attractive application areas for IoT. IoT technology has great potential for remote medical care (also named telemedicine or eHealth). It can effectively help doctors solve many problems caused by distance barriers. Sensor devices in the IoT can bring great convenience to telemedicine and have recently been developed for daily monitoring of physical activity awareness, personal health, and medical care. The smart wearable sensor devices used in remote health monitoring systems [1] collect and store a large amount of patient health status data (such as body temperature, blood pressure, and electromyography) through sensors. Afterwards, doctors can remotely monitor the health of patients in real time and make appropriate treatment recommendations accordingly until the patients have fully recovered. Hence, IoT technology is helpful to achieve an early diagnosis of patients and thus improving their health.

World Health Organization (WHO) [2] demonstrate that cardiovascular diseases (CVDs) are the

number 1 cause of death all over the world. Timely medical treatment at the time of the onset is of vital importance for patients with cardiovascular diseases. Thus, it has attracted increasing attention to telemedicine services based on IoT, which provide real-time medical care for interventional cardiovascular diseases and even make remote surgery possible. Stents are widely adopted to treat severe narrowing of the arterial lumen through implantation in patients. In clinical practice, it is essential to track the location of planted stents in a biomedical image, which determines the success or failure of stent surgery to some extent. But it is a great challenge as well. Traditionally, the task is completed by the arterial experts manually, requiring exhausting work and is time-consuming. As widely applied in biomedicine, deep learning [3] can realize the automation of stent detection. Therefore, combining deep learning with IoT enables a new direction towards the stent struct segmentation for cardiovascular diagnosis and treatment.

In this article, we introduce an eHealth framework based on IoT for interventional cardiovascular diagnosis and treatment. The framework provides an autonomous solution for cardiovascular patients who need interventional therapy and health monitoring. It mainly relies on wearable sensors that record information about the patient's health, such as respiratory rate, heart rate, and body temperature. Moreover, the framework combines the cloud-edge paradigm with deep learning to perform the stent struct segmentation to overcome the challenges of limited and imbalanced cardiovascular images. The local base classifier is deployed in the edge server to classify the local cardiovascular images. The pseudo-label generated from all the predicted results together with global data is fed to the image segmentation network in the cloud after data augmentation. Finally, asynchronous visualization telesurgery is adopted when the patient requires interventional therapy. The local doctors near the patient and the remote professional doctors interact with each other in real time to make surgical decisions. After the telesurgery, the framework restarts to the first step, and the remote doctors will continue to do the daily monitoring work. The entire framework is a cyclical process until the patient has fully recovered.

## LITERATURE REVIEW

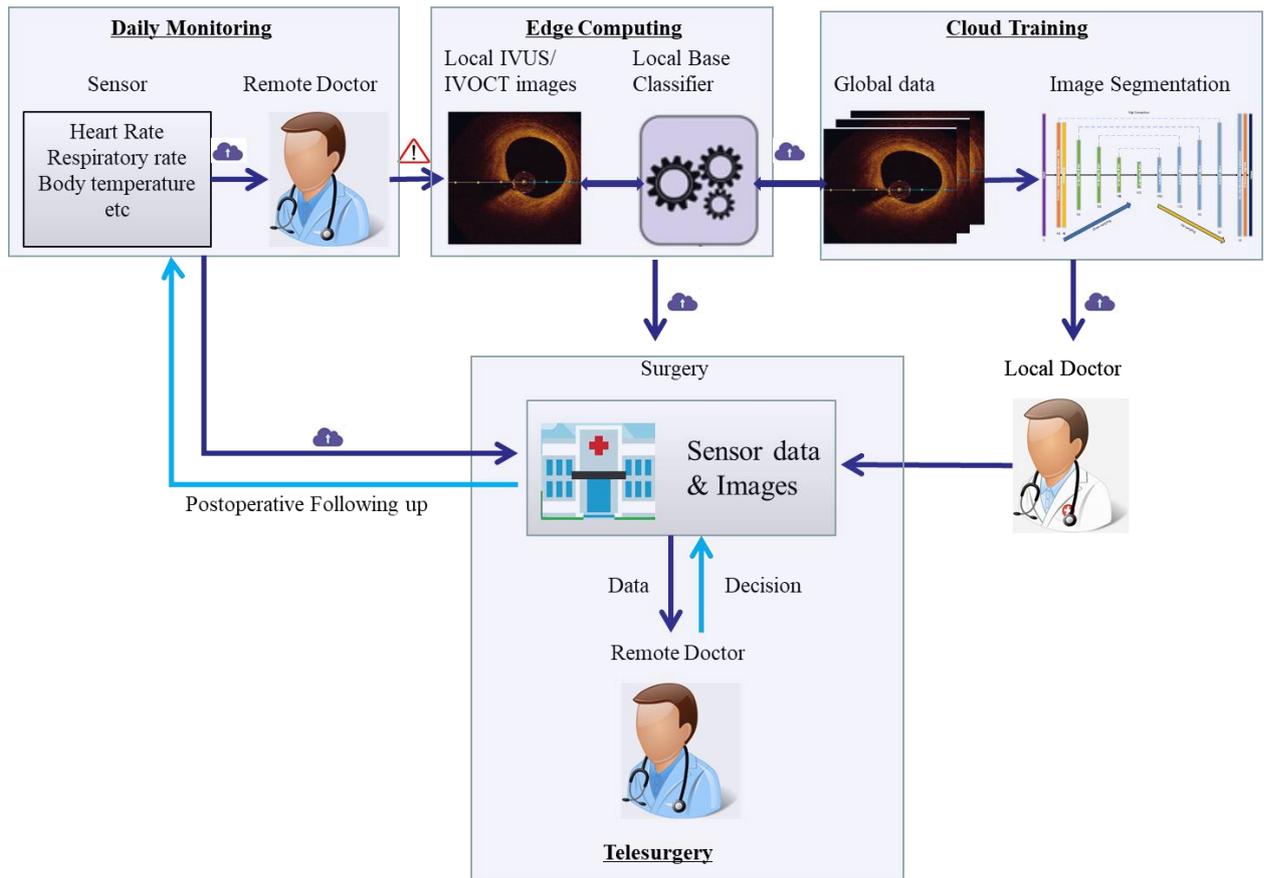
**IoT Architecture for Healthcare Systems:** In recent years, the potential viability and widespread use of electronic healthcare systems have set off a revolution in the field of healthcare. The aging of the global population and the increase in the number of patients are the two main factors driving this revolution. Moreover, the birth of IoT technology has brought a huge impact on the health care system. The decision support system based on the IoT is a typical part of recent healthcare [4]. Through this system, a patient's health is recorded to gain a deeper understanding of the patient's status for clinical decisions. Chatterjee *et al.* [4] ascertained the risk groups by embedding the logic of a Framingham score into the reasoning engine, where an input of a patient with all the parameters considered would return the risk score. However, they did not propose an efficient method of data collection in their research, and the proposed analysis system for running data is non-automated. With the advent of high-precision sensors and medical devices in the IoT, they can be regarded as smart devices or objects that constitute the core part of the IoT. The IoT devices are able to connect patients, clinics, and healthcare organizations seamlessly and securely via the Internet, making IoT based eHealth system a research trend. For patients with cardiovascular disease, doctors usually make a diagnosis based on certain parameters such as Electrocardiogram (ECG) signal, blood pressure, body temperature, and heart beat rate (HBR). Many scholars have also carried out researches on ECG monitoring based on the IoT. Most recently, Maity and Misra [5] proposed a more comprehensive patient data collection system by using specified sensors (HBR, ECG, and Heart sound), which can be applied in the remote monitoring and prevention of cardiovascular

diseases. However, there is a lack of an eHealth system for the diagnosis and treatment of interventional cardiovascular diseases so far. Thus, we propose an IoT framework for telemedicine that combines edge computing and cloud training.

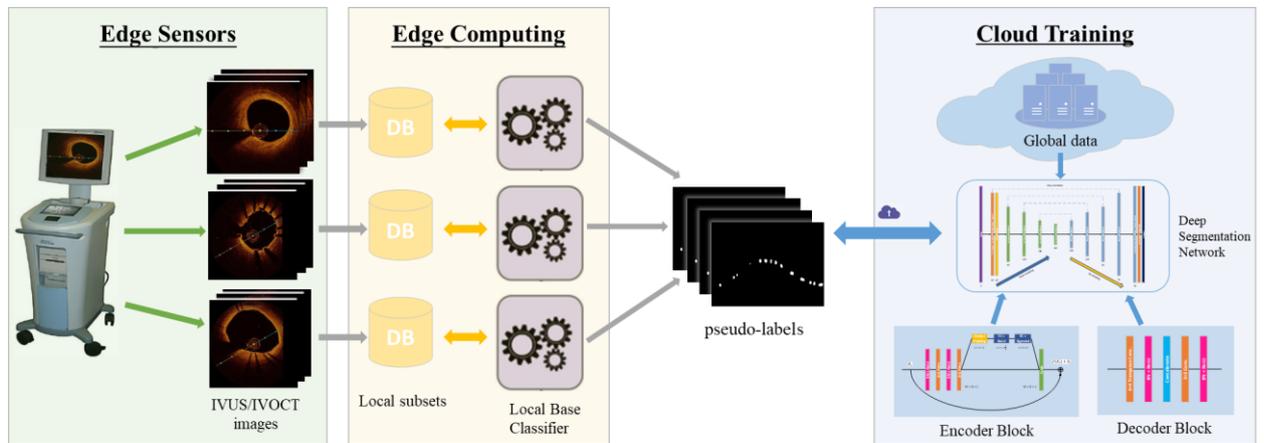
**Deep Learning for Stent Segmentation:** Bioresorbable Vascular Scaffolds (BVS), one of the most frequently used types of stents in cardiovascular interventional therapy, is absorbable and harmless compared to the metallic stent. In vascular medical imaging (e.g., intravascular optical coherence tomography (IVOCT), intravenous ultrasound (IVUS), and optical coherence tomography (OCT)), it effectively tracks the position of BVS. Currently, such a task is still mainly performed by the experts manually, which results in time-consuming and labor-intensive operations. In recent years, many researchers have realized the segmentation of BVS through machine learning. Although machine learning segmentation methods based on traditional features can almost achieve excellent results in stent segmentation tasks, these methods often depend on manually designed features, and its subjectivity may lead to large fluctuations in the performance of the constructed models. Recently, with the rapid development of deep learning, its end-to-end training pattern has effectively solved the defect. In terms of BVS segmentation, Cao *et al.* [3] proposed a robust BSV struts detection model based on Region-based Fully Convolutional Network (R-FCN), which used a Region Proposal Network (RPN) to extract struts region of interest (ROIs) in IVOCT images, and then classify the ROIs to realize the detection. To deal with limited training samples, they adopted transfer learning to train the detection model but there was a lack of data augmentation in their work. U-shaped convolutional neural network (U-net), a classical network widely used in medical image processing, has also been applied in BVS struts detection and segmentation by Zhou *et al.* [6]. Their network consisted of four up-sampling modules and five down-sampling modules. In this article, we enlarge our dataset through data enhancement and then propose a U-net based network with the pretrained SEResNet34 as its encoder. The proposed network outperforms the existing deep learning models and achieves a high segmentation accuracy, which will be able to aid the clinical surgery.

## The framework for Interventional Cardiovascular Diagnosis and Treatment

**An Overview of the Framework:** As illustrated in Fig. 1 (a), the framework can be divided into the following five parts. First, the remote doctor daily monitors the patient's data collected by the wearable sensor devices. When the heart rate is over 100/min, or diastolic blood pressure is over 90, or systolic blood pressure is over 140, the patient's physical signs can be regarded as abnormal. Then, the remote doctor will immediately inform the patient and ask him/her to go to the local hospital to obtain the IVUS/IVOCT images. After the image acquisition, the local training images are classified by the local base classifier in the edge server, and pseudo-labels will be generated by the output of the local base classifier. Then, after the data augmentation of the cloud global data and pseudo-labels, cloud training is performed on the deep segmentation network of stent struts. Next, the local and remote doctors complete telesurgery through real-time interaction and exchange of decisions. Finally, patients continue to wear sensor devices to record physical health data after telesurgery, and the remote doctor continues to monitor the data. The entire framework is a cyclical process until the patient has fully recovered.



(a)

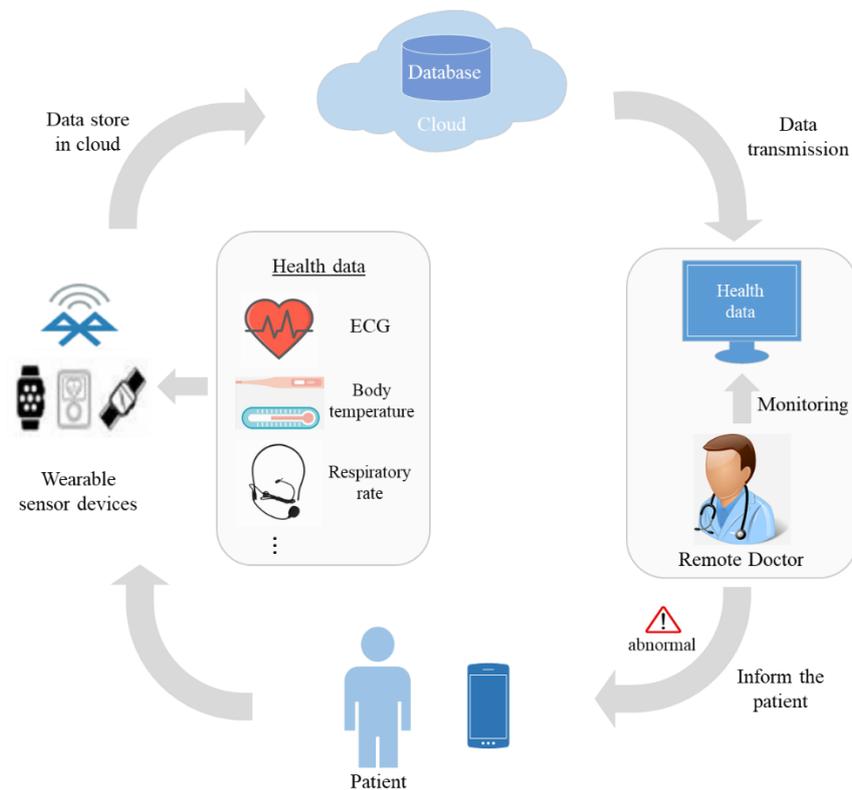


(b)

**FIGURE 1.** (a) The flowchart of the proposed framework for interventional cardiovascular diagnosis. The framework is divided into five parts: Daily Monitoring, Edge Computing, Cloud Training, Telesurgery, and Postoperative Following up. (b) Edge computing structure for IoT deep learning. The local base classifiers are deployed in edge nodes to train local images. The deep segmentation network is deployed in the cloud to train global data.

**Daily Monitoring:** In recent years, more and more scholars have tried to develop some smart wearable sensor devices [1] that can be used in remote health monitoring systems to continuously monitor the health of individuals. There is already a low-cost wrist-worn wearable sensor E-Vital which has an acceptable signal-to-noise ratio (SNR) and has achieved success in continuous data transmission, signal quality, and data storage performance [1]. The wrist-worn wearable sensor integrates several

sensors (such as accelerometer, temperature, ECG, and gyro) to monitor the patient's physical signs (such as body temperature and heart rate), real-time activity, and emotional state. In the first part of this framework (Fig. 2), wearable sensor devices are used to collect a large amount of patient's health data (such as respiratory rate, heart rate, and body temperature), which are archived and processed by connecting to the cloud services through the Internet. Then, the data is transmitted to the remote doctor via the cloud. The remote doctor constantly monitors the patient's health in real time based on these data and can quickly discover the patient's risky conditions. When the remote doctor observes an abnormality in the data, the remote doctor would notify the patient and collect images of the patient's cardiovascular internal conditions.



**FIGURE 2.** The process of daily monitoring. In this part, the remote doctor is constantly monitoring the health data collected from wearable sensor devices. When the data is abnormal, the remote doctor will inform the patient.

**Role of Edge Computing:** Currently, IVUS and IVOCT images are used to observe the internal conditions of coronary arteries. These two image acquisition techniques can provide more detailed information about the internal vessels for interventional cardiovascular diagnosis and treatment. As illustrated in Fig. 1 (b), the framework collects local IVUS/IVOCT images of patients through edge sensors deployed in edge nodes. These images are mainly divided into two datasets with one labeled and another not. Moreover, we introduce a deep segmentation network in the cloud that can work on global data without consuming excessive local resources. The local IVUS/IVOCT images are used to train the local base classifier located on the edge node. To make full use of the unlabeled dataset, we generate pseudo-tags by using the predicted results of the local base classifier with the average method as the decision function. Pseudo-tags are transmitted to the cloud services and used with real tags for cloud training, which reduces network traffic and pressure.

**Training at Cloud Server:** In the cloud, global data is used to train the deep segmentation network. Due to the limited training data, it is necessary to enhance the image data. For the deep segmentation network in the cloud, U-net [7] is adopted to realize stent struts segmentation. As shown in Fig. 3 (a), the network consists of a down-sampling path and an up-sampling path with 16 encoder blocks and 5 decoder blocks. The former path extracts feature from the input global data beginning with a  $7 \times 7$  convolution, a batch normalization (BN), and a ReLU activation. Then, a max-pooling layer is inserted before 16 encoder blocks. The encoder blocks are placed in 4 layers with 3, 4, 6, 3 ones in each layer. In the up-sampling path, 5 decoder units are placed and the first four of which are concatenated by skip connections with their corresponding layers in the contracting path as illustrated in Fig. 3 (a). In this way, both low-level details and high-level semantic information can be learned in this symmetrical architecture. At last, we use a  $3 \times 3$  convolution followed by a sigmoid activation layer to generate the output. The obtained segmentation result of stent struts is then transmitted to the local doctor still via the cloud.

To explain our network architecture in more detail, we present the internal structure of the encoder block and decoder block in Fig. 3 (b). The encoder block is constructed through integrating the Squeeze-and-Excitation Networks (SENet) [8] with residual connection [9].

In Fig. 3 (c), we provide a theoretical illustration of the Squeeze-and-Excitation (SE) mechanism. The SE block can be applied to any mapping and here we take convolution as an example. Since the convolution only takes place in a local space, it is difficult to obtain enough information to extract the relationship among channels. Hence, SENet proposed the Squeeze operation which encodes the entire spatial characteristics on a channel into a global feature through global average pooling, which is depicted in Fig. 3 (b). Therefore, the initial input is converted to the output which indicates the numerical distribution of the feature maps in this certain layer or global information in other words.

Next, we conduct the Excitation operation to extract the relationships among channels with the gating mechanism in the form of sigmoid. In this step, we adopt two full connection (FC) layer contained bottlenecks to reduce the complexity of the model and improve generalization performance. As can be seen in Fig. 3 (b), the former FC layer reduces the dimension with a ReLU activation followed, while the latter one restores the original dimension followed by a sigmoid activation. Finally, we multiply the initial characteristics with the learnt activation values of each channel. This is the corresponding step ‘scale’ drawn in Fig. 3 (b). In general, the whole Squeeze-and-Excitation operation can be regarded as a learning process of each channel’s weight coefficient, making the model more discriminative to the characteristics of different channels.

In our implementation, the SE module is embedded in the residual learning branch in the encoder block shown in Fig. 3 (b). As a ResNet unit, if we denote the input of each encoder block as  $X_i$ , the output will be  $f(X_i) + X_i$  with  $f(\cdot)$  representing the residual function. During the encoding process, the down-sampling block begins with two sets of BN + ReLU, and each one is followed by a  $3 \times 3$  convolution layer. The SE block is inserted right after the latter convolution. As for the decoder block, it is constructed by 5 layers: a  $4 \times 4$  transpose convolution, BN + ReLU, Concatenate,  $3 \times 3$  convolution, and BN + ReLU in turn. Note that a skip connection is linked before each Concatenate layer in the first four decoder blocks, connecting different levels of information to improve the deep segmentation network.

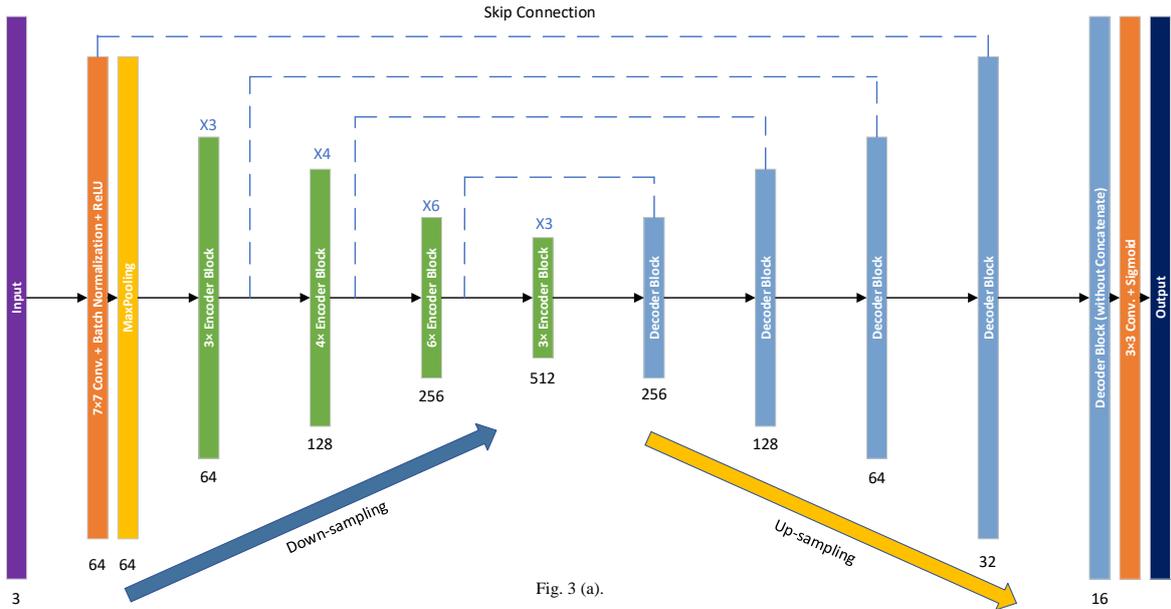
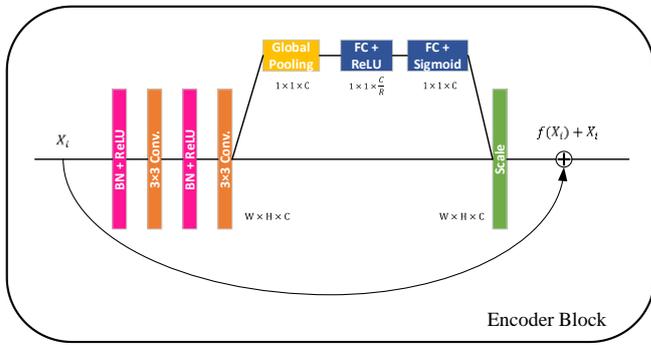
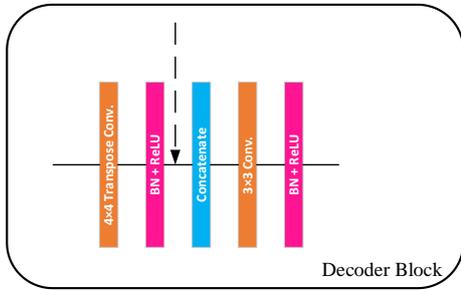


Fig. 3 (a).



Encoder Block



Decoder Block

Fig. 3 (b).

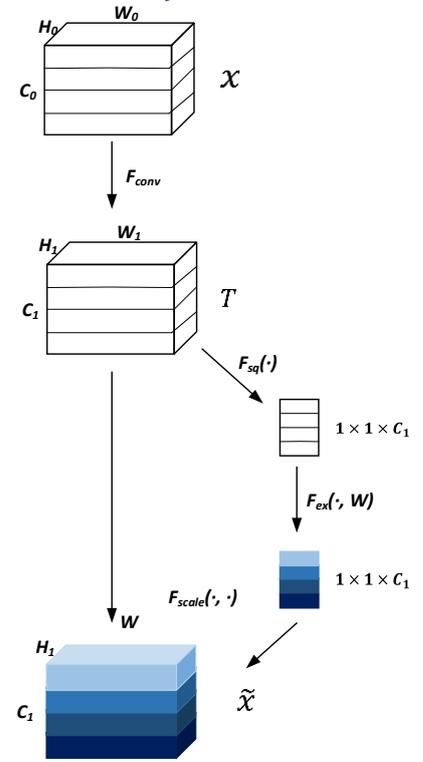


Fig. 3 (c).

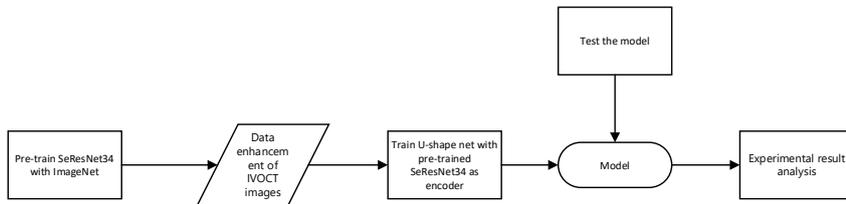


Fig. 3 (d).

**FIGURE 3.** (a)The architecture of the deep segmentation network. (b) The internal structure of the encoder block and decoder block. (c) The Squeeze-and-Excitation mechanism. (d) The training and testing process.

Due to the limited training data and even fewer stent contained training images, there is a strong imbalance between positive and negative samples. Therefore, Dice Loss (DL) is used as the loss function of the network. Additionally, we adopt adaptive moment estimation (Adam) Optimizer to train our network. The initial learning rate is set to 0.001 and decays by 0.15. We set the batch size to 4 and a total of 25 epochs are trained.

**Telesurgery:** If the patient's condition is serious to a level where surgical treatment is required, then a remote surgery will be performed. Firstly, the sensor data and images are transmitted to the local doctor and remote doctor via the cloud as the basis for diagnosis. During the operation, real-time interaction and visualization are achieved between the doctors on both sides through cloud services. Thus, the two sides can share their decisions instantly, enabling specialized surgeons to assist and guide doctors with less experience remotely. As the technical support for telesurgery, the transmission quality and speed of the image/video communication must be guaranteed. However, visual distortions after transmission caused by limited equipment conditions may interfere with the operation and even endanger patients' health. In the area of eHealth, there already exist some medical image quality evaluations such as Double Stimulus Continuous Quality Scale (DSCQS) and Mean Opinion Score (MOS). We adopt these metrics to further measure and improve the perceived quality of telecommunication.

**Postoperative Following up:** After the remote surgery, patients will continue to wear sensors so that their health data can be continuously recorded and sent to the remote doctor via the cloud. The remote doctor is responsible for follow-up health monitoring just as with their routine work. Actually, the entire framework forms a cyclical structure applying to the treatment of each patient until he/she has fully recovered.

## Experiment

**Experiment Set:** We collect the IVOCT images dataset of bioabsorbable stents from the department of cardiology of Dongfang hospital affiliated to Tongji University. A total of 1040 labeled photos are provided from experts, among which 624 images are used as training set while 208 ones as testing and validation set. The network trains on the training set and validates on the validation set, using the validation error as the stopping criteria of the training process. Our experimental results are derived from the performance on the testing set that is unused during the training process. Besides, we adopt data augmentation methods, such as random rotation with 0 – 90 degree, width and height shifting with range 0.2, vertical flipping, and horizontal mirroring, to enhance the training set to avoid overfitting. The training set has been expanded to three times the original set.

**Metrics:** Accuracy, sensitivity, Jaccard similarity coefficient, and Dice similarity coefficient are four commonly used evaluation criteria for medical image segmentation. They range from 0 to 1, and the closer to 1, the predicted contour is more similar to the ground truth.

**Result:** The training and testing process of our experiment is described in Fig. 3 (d). Firstly, we adopt transfer learning through pre-training the SEResNet34 with ImageNet, which can simplify the learning process and improve the segmentation effectiveness at the same time. As for the preprocessing, data enhancement of the IVOCT images is conducted to tackle the limited data. After comparison, the pretrained SEResNet34 is used as the encoder of the U-shape network. Finally, the trained model is validated by the testing set. We conduct the experiment based on the Keras framework with a GPU of

NVIDIA GeForce GTX 1080 (8GB).

To determine the structure of the encoder and decoder, we perform two comparative experiments respectively for selection. As listed in Table 1, we find that a transferred network performs better in most cases. Also, ResNet added with Squeeze-and-Excitation (SE) modules shows improved effectiveness than its original model because the SE mechanism enables the network to focus the attention on the important channels. Since a deep network is able to extract more features, ResNet34 performs better than ResNet18. However, a deeper network, such as ResNet50, will cause the loss of information and unclear segmentation boundaries due to a limited number of images in our dataset. Thus, we choose SEResNet34 to be the encoder due to its best results among the comparative tests. In addition, we record the parameters of these encoders in Table 1. Both the model size and prediction time are not at a large level for the limited training data. Nevertheless, the gap can be widened when a huge amount of data is accessible in future practical applications.

Encoder	Decoder	Transferable	Accuracy (%)	Sensitivity (%)	Jaccard (%)	Dice (%)	Model size (MB)	Prediction time (second / image)	Epochs to converge
ResNet18	transpose convolution + Batch Normalization	N	89.4	91.0	88.3	91.0	172	0.28	15
		Y	89.7	91.1	88.4	92.1			15
ResNet34	transpose convolution + Batch Normalization	N	90.3	90.7	89.8	92.1	294	0.75	15
		Y	90.5	90.7	90.0	92.5			20
SeResNet 18	transpose convolution + Batch Normalization	N	90.5	90.8	90.3	92.4	174	0.80	17
		Y	90.6	91.1	90.4	92.6			20
SeResNet 34	upsampling	Y	90.4	91.0	89.7	92.8	311	0.92	25
	transpose convolution	Y	90.9	91.3	90.0	93.0			25
	upsampling + Batch Normalization	Y	90.8	91.5	90.2	92.9			25
	transpose convolution + Batch Normalization	N	90.8	91.5	90.7	93.0			25
	<b>transpose convolution + Batch Normalization</b>	<b>Y</b>	<b>91.1</b>	<b>91.9</b>	<b>90.7</b>	<b>93.3</b>			<b>30</b>
SeResNet 50	transpose convolution +	N	91.0	91.3	90.5	92.8	422	1.31	40
		Y	90.9	91.1	90.4	92.9			50

	Batch Normalization				
--	---------------------	--	--	--	--

TABLE 1. The comparison of network structures.

When it comes to the decoder, we fix the pretrained SEResNe34 as our encoder for subsequent operations. When used alone, a transpose convolution works better than an upsampling in all the four metrics, and it is the same with their effects improved by adding Batch Normalization (BN). Therefore, the final decoder is constructed by a transpose convolution plus with BN.

Finally, we demonstrate the reliability of our deep segmentation network by comparing our results with other state-of-art methods on the same dataset. The segmentation outputs are given in Fig. 4 and the quantitative results are summarized in Table 2.

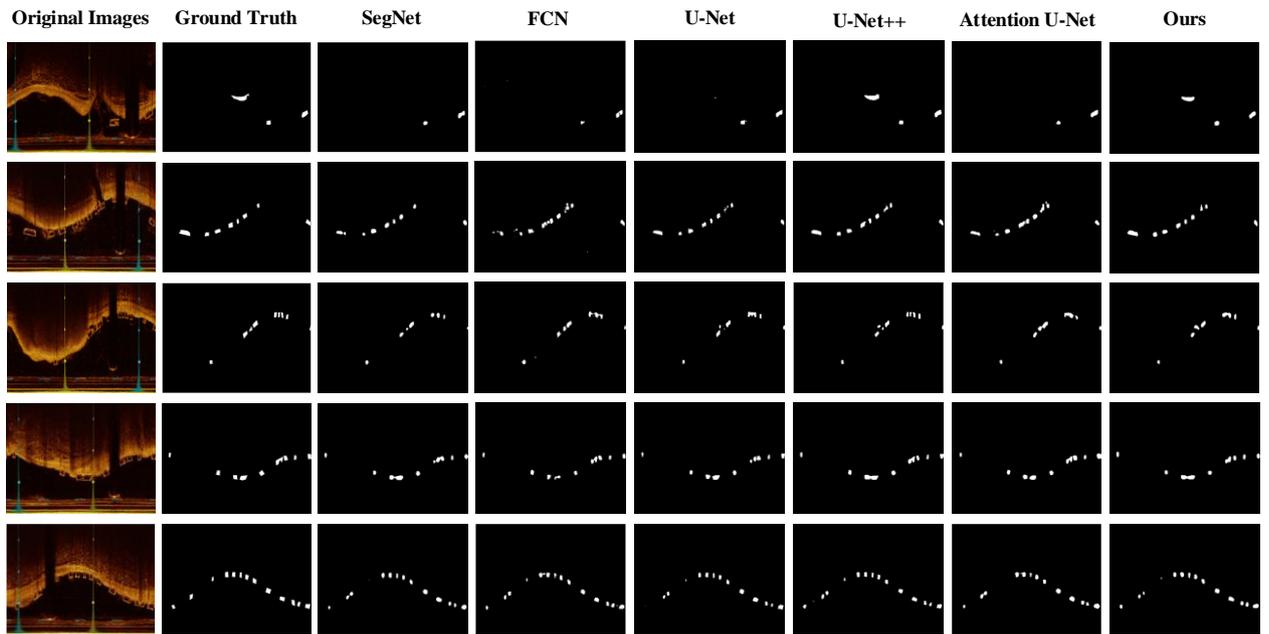


FIGURE 4. The segmentation outputs of images.

Model	Accuracy (%)	Sensitivity (%)	Jaccard (%)	Dice (%)
SegNet[10]	86.9	89.3	88.6	91.2
FCN[11]	87.1	90.7	87.9	92.1
U-Net[7]	87.6	90.3	86.3	91.5
U-Net++[12]	89.3	91.4	88.1	91.7
VGG_UNet[13]	89.3	91.4	88.1	91.7
Resnet_Unet[14]	89.9	91.8	89.2	92.3
Attention U-net [15]	90.1	91.3	89.4	92.5
<b>Ours</b>	<b>91.1</b>	<b>91.9</b>	<b>90.7</b>	<b>93.3</b>

TABLE 2. The comparison of other state-of-art methods.

## CONCLUSIONS

In this article, we introduce an IoT-based eHealth system for interventional cardiovascular disease. The system can effectively monitor the health of patients in real time through sensors and provide timely interventional telesurgery treatment for patients when necessary. Besides, we introduce the deep learning

of the Internet of Things into the edge computing environment, reducing the network pressure between IoT devices and cloud services. In the experiments, we choose U-net as the backbone for the deep segmentation network and use the pretrained SEResNet34 as its encoder. Extensive experiments have been conducted to prove the effectiveness and robustness of the proposed method, and accuracy, sensitivity, jaccard, and dice of 91.1 percent, 91.9 percent, 90.7 percent, 93.3 percent are achieved respectively. In the future, various deep learning methods can be introduced to propose an adaptive deep learning network for different kinds of data images from different distribution. Moreover, high latency and reliability remain the focuses of future implementation as required by the medical industry. With the development of the 5th generation wireless systems (5G), while guaranteeing end-to-end millisecond-level latency and nearly 100 percent reliability, it will be possible to deal with massive data collected from every edge of the world.

## ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (Grant No. 2020YFB1313703), the National Natural Science Foundation of China (Grant No. 62002304), and the Natural Science Foundation of Fujian Province of China (Grant No. 2020J05002).

## REFERENCES

- [1] F. P. Akbulut, Ö. Özgür, and İ. Çınar, "e-Vital: A Wrist-Worn Wearable Sensor Device for Measuring Vital Parameters," *2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey*, 2019, pp. 1-4.
- [2] World Health Organization, "Cardiovascular Diseases," <https://www.who.int/health-topics/cardiovascular-diseases>, accessed August 31, 2020.
- [3] Y. Cao *et al.*, "Deep Learning Based Bioresorbable Vascular Scaffolds Detection in IVOCT Images," *2018 24th International Conference on Pattern Recognition (ICPR), Beijing*, 2018, pp. 3778-3783.
- [4] P. Chatterjee, L. J. Cymberknop, and R. L. Armentano, "IoT-based decision support system for intelligent healthcare – applied to cardiovascular diseases," *2017 7th International Conference on Communication Systems and Network Technologies (CSNT), Nagpur*, 2017, pp. 362-366.
- [5] A. Maity and I. S. Misra, "Prototype Design of An IoT Enabled Cost-Efficient Portable Heart-Health Data Acquisition System," *2020 IEEE Calcutta Conference (CALCON), Kolkata, India*, 2020, pp. 137-141.
- [6] W. Zhou *et al.*, "Automatic Detection Approach for Bioresorbable Vascular Scaffolds Using a U-Shaped Convolutional Neural Network," in *IEEE Access*, vol. 7, pp. 94424-94430, 2019, doi: 10.1109/ACCESS.2019.2926523.
- [7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *MICCAI*, 2015, pp. 234–241.
- [8] Hu J *et al.* "Squeeze-and-Excitation Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, PP(99).
- [9] K. He *et al.*, "Deep residual learning for image recognition," *CVPR*, 2016, pp. 770–778.
- [10] Badrinarayanan V, Kendall A, and Cipolla R. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, 2017, 39(12): 2481-2495.
- [11] Chen L C *et al.*, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," arXiv preprint arXiv:1412.7062, 2014.
- [12] Zhou Z *et al.*, "Unet++: A nested u-net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, Cham, 2018: 3-11.
- [13] Iglovikov V and Shvets A, "TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation," arXiv preprint arXiv:1801.05746, 2018.
- [14] Xiao X *et al.*, "Weighted Res-UNet for High-Quality Retina Vessel Segmentation," *2018 9th International Conference*

on *Information Technology in Medicine and Education (ITME)*, *IEEE*, 2018: 327-331.

[15] Oktay O *et al.*, "Attention u-net: Learning where to look for the pancreas," arXiv preprint arXiv:1804.03999, 2018.

## BIOGRAPHIES

Chenxi Huang (supermonkeyxi@xmu.edu.cn) received his Ph.D. degree from Tongji University, Shanghai, China, in 2019. He is currently an assistant professor with Xiamen University, Fujian, China. His research interests include deep learning, multimedia analysis and medical image processing.

Yongshuo Zong (1750774@tongji.edu.cn) is currently pursuing the B.S. degree with Tongji University, Shanghai, China. His research interests include medical image analysis and deep learning.

Jinling Chen (24320172203113@stu.xmu.edu.cn) is currently pursuing the B.E. degree with Xiamen University, Fujian, China. Her research interests include deep learning and medical image processing.

Weipeng Liu (liuweipeng@hebut.edu.cn) received his Ph.D. degree in Hebei University of Technology. Currently, he works as a professor in Hebei University of Technology as well. His research interests include robot vision, artificial intelligence, and deep learning. He has published more than 20 papers and received several awards in robot control.

Prof. Jaime Lloret (jlloret@dcom.upv.es) [M'07–SM'10] received his B.Sc.+M.Sc. in Physics in 1997, his B.Sc.+M.Sc. in electronic Engineering in 2003 and his Ph.D. in telecommunication engineering (Dr. Ing.) in 2006. He is an associate professor at the Universitat Politècnica de València, Chair of the Research Institute IGIC, and the Head of the Active and Collaborative Techniques and Use of Technologic Resources in the Education (EITACURTE) Innovation Group. He was the Vice-Chair for the Europe/Africa Region of the Cognitive Networks Technical Committee of the IEEE Communications Society for the term 2010–2012 and Vice-Chair of the Internet Technical Committee of the IEEE Communications Society and the Internet Society for the term 2011–2013. He is also the Chair of the Working Group of the Standard IEEE 1907.1. Since 2016, he has been the Spanish Researcher with the highest h-index in the telecommunications journal list according to the Clarivate Analytics Ranking. He is an ACM Senior Member and IARIA Fellow.

Mithun Mukherjee (m.mukherjee@ieee.org) received the B.E. degree in electronics and communication engineering from the University Institute of Technology, Burdwan University, Bardhaman, India, in 2007; the M.E. degree in information and communication engineering from the Indian Institute of Science and Technology, Shibpur, India, in 2009; and the Ph.D. degree in electrical engineering from the Indian Institute of Technology Patna, Patna, India, in 2015. Currently, he is a professor with the school of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China. His research interests include fog computing and intelligent edge computing.