IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 15, 2019, accepted December 4, 2019, date of publication December 10, 2019,
date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958825

Diagnostic Model of Coronary Microvascular Disease Combined With Full Convolution Deep Network With Balanced Cross-Entropy Cost Function

SHIWEN PAN, WEI ZHANG, WANJUN ZHANG, LIANG XU, GUOHUA FAN, JIANPING GONG, BO ZHANG, AND HAIBO GU
The Second Affiliated Hospital of Soochow University, Suzhou 215000, China

Corresponding authors: Bo Zhang (bozhang2019@aliyun.com) and Haibo Gu (haibogu@aliyun.com)

This work was supported in part by the Social Development Guiding Program of Suzhou City in China under Grant SYSD2019110, in part by the Radiomics of Multiparametric MRI Associated with Treatment Responses and Survival Outcomes of Lung Cancer Brain Metastases after Whole Brain Radiotherapy, and in part by the Pre-Research Fund Program of the Second Affiliated Hospital of Soochow University under Grant SDFEYQN1205.

ABSTRACT This paper addressed the vessel segmentation and disease diagnostic in coronary angiography image and proposed an Encoder-Decoder architecture of deep learning with End-to-End model, where Encoder is based on ResNet, and the deep features are exacted automatically, and the Decoder produces the segmentation result by balanced cross-entropy cost function. Furthermore, batch normalization is employed to decrease the gradient vanishing in the training process, so as to reduce the difficulty of training the deep neural network. The experiment results show that the algorithm effectively exacts the feature and edge information, therefore the complex background disturbance is suppressed convincingly, and the vessel segmentation precision is improved effectively, the segmentation precision for three typical vessels are 0.8365, 0.8924 and 0.6297 respectively; and the F-measure are 0.8514, 0.8786 and 0.7298, respectively. In addition, the experiment results show that our proposed can be generalized to the angiography image within limits.

INDEX TERMS Coronary microvascular, cross-entropy, cost function, encoder-decoder, deep learning, batch normalization.

I. INTRODUCTION

Cardiovascular disease is very common and one of the diseases with the highest morbidity in the world [1]. With the development of medical imaging technology, doctors can have a intuitive understanding of the cardiovascular disease through medical imaging. Nowadays, many imaging techniques for the diagnosis and treatment of coronary heart disease have been developed. Computed Tomography Angiography (CTA) has the advantages of low price, simple operation and non-invasive imaging and is one of the most commonly used diagnostic methods for coronary heart disease [2].

With the development of the economy and the increase of life pressure, the prevalence of cardiovascular disease is rising, it is one of the most common causes of death in the world, which is a serious threat to human health [3]. Studies have shown that coronary vascular stenosis is the main cause of coronary heart disease. In the human body, the coronary artery is the only way to supply blood to the cardiac muscle, and its health determines whether the cardiac muscle has enough nutrients/energy to maintain the rhythmic beat of the heart. Stenosis of the coronary arteries usually causes stenosis or obstruction of the vascular lumen by certain precipitates in blood vessels or chronic thickening of the vessel wall, which impedes the normal flow of blood and leads to myocardial ischemia or necrosis [4]. Figure 1 is myocardial ischemia caused by coronary atherosclerosis. There are many diagnostic methods for coronary vascular stenosis. Coronary angiography is a safe, reliable and effective non-invasive diagnostic method, as shown in Figure 2. It is still the gold standard for the diagnosis of coronary heart disease at home

The associate editor coordinating the review of this manuscript and approving it for publication was Yongtao Hao.

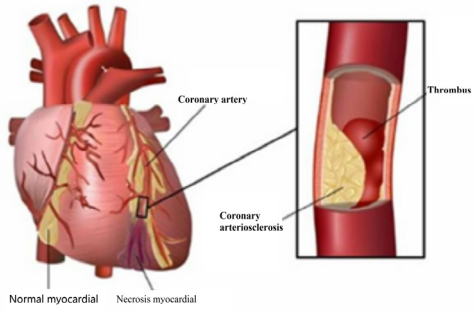


FIGURE 1. Myocardial ischemia caused by coronary atherosclerosis.

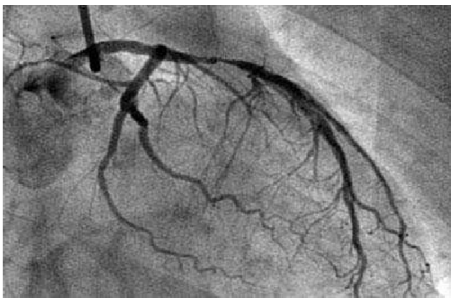


FIGURE 2. Coronary angiography.

and abroad. At present, the diagnosis method of coronary angiography image is that the doctor analyzes the blood vessel tree structure in the image, analyzes the contrast of the pixel, and determines the disease condition based on experience [5]. Due to the limitation of imaging quality and personal experience, the results of manual reading are susceptible to subjective factors, and have the disadvantages of time-consuming and labor-intensive, low reliability and poor repeatability.

With the development of computer technology, medical image processing and analysis technology is increasingly used in clinical practice so as to achieve an auxiliary diagnosis of the disease, help doctors reduce repetitive work, improve diagnostic efficiency and accuracy, and develop optimal treatment options. Recently, the effective method for judging angiographic stenosis is to locate the location of stenosis based on the blood vessel diameter parameter in the image [6]. Therefore, the stenosis detection and recognition based on blood vessel diameter measurement is of great significance. Extracting the contour (or edge) of the blood vessel and the centerline of the blood vessel to calculate the diameter of the blood vessel is a common method for determining the stenosis of blood vessels. Methods for extracting blood vessel contours include mathematical morphology methods, parametric model methods, and statistical methods. The common methods for extracting centerlines are based on vascular region tracking and geometric feature methods. When calculating the diameter of blood vessels, the more widely used algorithms mainly include edge intersection method and parametric model method [7]. The edge intersection method is a relatively simple diameter measurement method, and the accuracy of the edge point position directly affects the diameter measurement result. The parametric model method estimates the matching degree between the model parameters and the actual blood vessel. In addition, Kalman filtering is used to detect areas of sudden changes in vessel diameter for analysis of vascular stenosis. This method treats the diameter of the blood vessel as a series of signals, and the abnormal region of the signal obtained by the analysis is the stenosis region of the blood vessel. The stenosis area detected by this method has a certain hysteresis in space. At present, most researchers focus on blood vessel contour extraction, centerline extraction, and diameter calculation [8].

Therefore, our works aim to provide a new model for clinical diagnosis by combining the deep learning technique and the balanced Cross-entropy cost function to achieve automatic segmentation and stenos is assessment in coronary angiography image. This paper builds Encoder-Decode framework based on deep residual network (ResNet), completes the automatic extraction of image features and the learning of segmentation models, thus achieving end-to-end segmentation result from input image to the segmentation result. In addition, the normalization processing of data is implemented by Batch Normalization technology before each convolution operation, thus discarding the traditional convergence strategies such as Dropout and Weight Decay, and improving the training precision of the neural network. Finally, a verification experiment was carried out for the blood vessels in the angiographic image database.

II. RELATED WORKS

The use of statistical region fusion method for blood vessel segmentation is currently the most common method. This method was first proposed by Nock et al. The main idea is to determine the similarity of adjacent pixels and regions in the image through statistical criteria so as to realize the integration. This method not only considers the difference in gray scale between different areas in the image, but also specifies the actual meaning of the corresponding label in different areas. For example, there are differences in gray value, contrast in an angiographic image, but they are all blood vessel region. Therefore, the method can accurately fuse these different regions into the same blood vessel region [9]. The segmentation method consists of three steps: (1) statistical description of the image region: the image is regarded as a series of statistical features; (2) statistical description of region fusion: regional fusion is done according to the statistical characteristics of the image. (3) regional fusion: judging all the pixels one by one according to the obtained criterion, finding the regions with the same attributes and merging them. Although statistical region fusion method performs well when extracting dim-small blood vessels, it is easy to cause false detection under CTA images.

With the continuous development of deep learning technology, it has shown outstanding advantages in the field of computer vision and image processing. More experts and scholars have begun to apply deep learning technology to

the field of vessel segmentation in angiogram. In [10], vessel segmentation based on deep learning combined with semi-supervised adaptive support vector machine is used to improve the segmentation accuracy in CTA images, and the vessel center line and the microvessel with low contrast are positioned by radial projection, and the improved complex wavelet is used to enhance the blood vessel. Deep learning is applied to the enhanced images to generate deep feature vectors, and then semi-supervised support vector machine is performed to extract the main structure of the blood vessels. Finally, the segmented image is a combination of tiny blood vessels and main blood vessels.

From the research of 2D sequence image segmentation, literature [11] proposes an improved particle filter algorithm based on multi-features, which realizes the tracking and segmentation of vascular in coronary CTA sequence images, and render the 3D model of vascular by surface rendering algorithm. In the tracking algorithm, firstly, the bifurcated coronary vessels are tracked by the feature matching algorithm; secondly, the multi-features are combined and the particle resampling rules are improved to realize the tracking of small vascular with topological structure and positional changes; Finally, the center of the tracking target is used as a seed point to achieve sequence segmentation of vascular [12]. Through experimental data analysis, the accuracy of coronary target tracking in this algorithm reaches 97.84%, and compared with other tracking methods, the superiority of the algorithm in this paper is verified.

In order to realize automatic segmentation, from the research of 3D volume data segmentation, this paper proposes a 3D W-Net with adaptive weight loss function (3D W-Net With Adaptive Weighted Loss [13], AWL-W-Net). In this process, the W-Net network structure training segmentation model is established; Secondly, combined with the characteristics of coronary CTA images and W-Net network training results, the adaptive weight loss function layer is proposed, which enhances the network learning ability and achieves the goal of improving the segmentation accuracy of the vascular model and repairing the fracture of the 3D model. AWL-W-Net can complete the high-precision segmentation, and compare the results with W-Net and 3D U-Net network, which proves that AWL-W-Net network is not only improved the accuracy of segmentation, and the fracture of the model can be repaired, providing doctors with a high-precision 3D model of coronary vessels that is more in line with clinical needs.

In summary, the vessel segmentation based on deep neural network in coronary angiography shows that the feature space construction is the process of extracting image features by deep neural network. The accuracy is determined by the structure of deep neural network, and it will also affect the accuracy of blood-vessel segmentation. In order to further approximate the strong nonlinearity of the segmentation model, the depth of the deep network is generally set deeper. This often leads to gradient vanishing or gradient explosion of error back propagation during network training, making deep neural networks difficult to train. Therefore, the key to achieving end-to-end high-precision segmentation of blood vessels is to construct a reasonable and well-trained deep learning framework and structure.

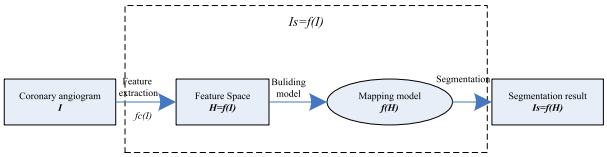


FIGURE 3. Flow diagram of vessel segmentation in angiographic image.

III. PROBLEM DESCRIPTION

Given the angiographic image I, the pixel point is $g_h \times g_w = g$, where g_h is the height of I, g_w is the width of I, and the number of information channels is C; if the image matrix is transformed into vector, there is $I = \{x_1, x_2, \cdots x_g\} \subset R^{g\times c}$; the blood vessel segmentation result of angiographic image [13] is $I_s = \{y_1, y_2, \cdots y_g\} \subset R^{g\times 2}$. And the segmentation process is shown in Figure 3.

The essence of vascular segmentation is to establish an end-to-end mapping relationship between the angiographic image I and the segmentation result I_s. Because of the high dimension of the input image I, the complex background and the complex and changeable vascular structure, generally $f(\cdot)$ is a high-dimensional nonlinear model, it is difficult to establish the mapping relationship $I_s = f(I), f : g \times c \rightarrow g \times 2$ at one time. Therefore, the process is divided into two steps: first, extract features and establish feature space; then, build the reasonable segmentation mapping model $f_H(\cdot)$ based on the feature space to complete the segmentation of blood vessels [13].

If the feature space of the segmentation model is set as $H = h_1, h_2, \cdots, h_n \subset R^{n\times m}$, there is a mapping relationship $f_c(\cdot), f_c : g \times c \rightarrow n \times m$ between I and H; then complete the feature extraction of the image, and there is $H = f_c(I)$ [15]. Obviously, the mapping model between the feature space and the segmentation result is $f_H(\cdot), f_H : n \times m \rightarrow g \times 2$, and the segmentation of blood vessel $I_s = f_H(H)$ is realized.

With the continuous improvement of resolution on angiographic images, there are two challenges in establishing the mapping model between image feature space and blood vessels:

(1) The dimension of the feature space has been greatly improved. For the same blood vessel, the improvement of resolution on the angiographic image means that the number of pixels representing the target will increase. If $I \subset R^{g\times c}$ only represents blood vessel, then G must increase with resolution. This will inevitably make the mapping relationship $I_s = f(I)$ between the angiographic image and the segmentation result show high-dimensional characteristics [16].

At the same time, with the improvement of resolution, the structure, texture, spectrum and other information of angiographic image are more abundant, and the number of features that can be used to effectively represent blood vessels is

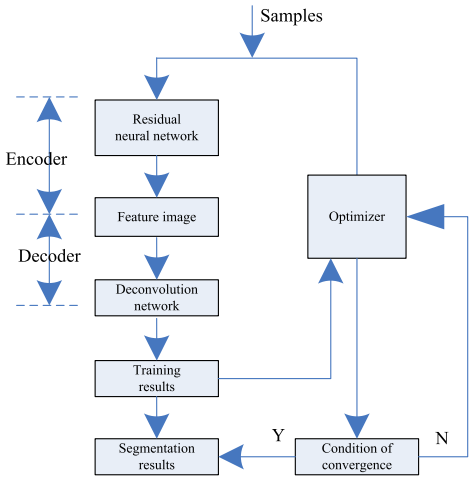


FIGURE 4. Deep learning framework.

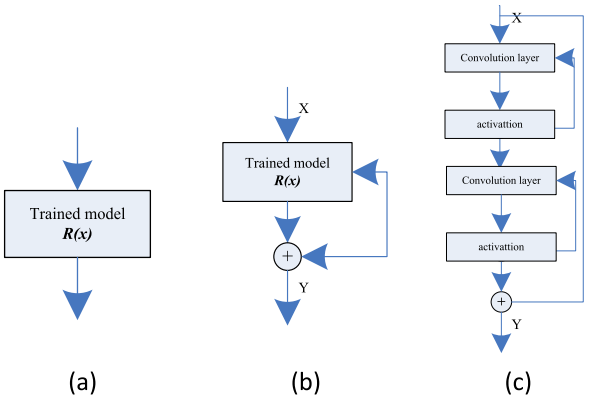


FIGURE 5. Deep learning framework; (a) Conventional framework; (b) Residual model framework (c) our proposed framework.

bound to increase, that is, the dimension of feature space H of angiographic image is greatly increased. So this inevitably makes it more difficult to build $f_c(\cdot)$ and $f_H(\cdot)$ manually. This is also one of the main reasons for the low accuracy, low generalization ability and even failure of traditional image segmentation methods [17].

(2) The non-linear relationship of the mapping model is more complex. In low-resolution angiographic images, a vessel may only be represented by a few or dozens of pixels, and the simple segmentation model $f_H(\cdot)$ can be directly established through specific spectral information (or gray-scale information). In high-resolution angiographic images, the dimension of feature space H increases, and feature variables (such as gray scale, information entropy, edge information, etc.) are often coupled with each other, and the relationship between feature variables and segmented targets is strong nonlinear. This will inevitably lead to the manual segmentation model $f_H(\cdot)$ cannot accurately describe the relationship between the image and the 0bject.

IV. OUR PROPOSED MODEL ARCHITECTURE

A. THE GENERAL FRAMEWORK

Deep learning is to train a large number of samples to make the trained deep neural network approach the real model $I_s = f(I)$ without intermediate process, so that the end-to-end task mode from the input image to the segmentation result can be realized [18]. In this paper, the high-resolution angiographic image segmentation based on deep learning framework is shown in Figure 4.

Each batch contains two parts: angiographic image and vascular label. The feature image (from which the feature space H is expanded) is obtained by making angiographic image go through the deep residual neural network, and the function of encoding vascular features is realized [19]. Then, a deconvolution decoder is used to segment the blood vessels through the up-sampling of the feature image, and the output size is restored to the original size of the angiographic image. And this process realizes the function of decoder. The decoding results are sent to the optimizer together with the vascular label, and the weights of the residual neural network and the deconvolution network are optimized by using the random gradient descent method [20]. When the training is over, the weight of the fixed neural network remains unchangeable, and the output result is the segmentation result of the blood vessel.

Residual neural network is a network structure proposed to reduce the difficulty of deep neural network training. The training object is no longer the real model $R(x)$, but the difference $R(x) - x$ between the real model and the input sample. The frame structure of residual neural network is shown in Figure 3 (b). where the trained model is residual model $F(x) = R(x) - x$, and the feed-forward channel of the sample is introduced at the output end of the model to form a closed loop, which makes $y = F(x) + x = R(x)$, so the final output is still the real model. In this framework, the weight convergence of neural network is more effective.

B. ENCODER STRUCTURE BASED ON RESIDUAL DEEP NEURAL NETWORK

Suppose the sample input is x; the training output is $y = R(x)$; and the conventional convolutional neural network is shown in Figure 5(a); the training result directly approximates the real model. In order to make the output model more approximate to the real model, the deep learning network is usually designed deeply to obtain high-dimensional and strong non-linear mapping. But it often makes the network training difficult, which leads to the decrease of prediction accuracy and even training failure. For this reason, He et al. proposed a residual neural network [21], which can better approximate the real model of the system, and its principle is shown in Figure 5(b).

The implementation framework of single-layer residual neural network is shown in Figure 5(c). The activation layers are after the front-end convolution layer, and the output of the last convolution layer and the feed-forward signal are superposed to activate the output. In the closed loop of the feed-forward channel, there are many convolution layers and activation layers, but in this paper, a typical two-layer convolution is used.

Under the guidance of the core idea of Literature [23], this paper uses the deep residual neural network as the basis to build the encoder structure to automatically extract the features needed for vascular segmentation. However, the level of the typical residual neural network is generally designed deeper, and usually there are 50, 101, 200 or more layers. The deeper the network structure is, the more computation will be required. After a lot of experiments, on the premise of ensuring the segmentation accuracy, a 31 layer residual neural network is designed to reduce the calculation. The deep residual network is divided into five convolution types, and the number of each convolution layer is set to 1, 6, 8, 8, 8, respectively. The number of convolutions can not only accurately extract the features needed for vascular segmentation, but also significantly reduce the amount of calculation. Except for the convolution layer of the first type, the difference of other convolution types mainly lies in the number of convolution kernels.

In literature [24], the residual model framework is described, where K is the convolution kernel size; S is the convolution step; and C is the number of convolution output channels. Each convolution layer includes convolution, Relu activation and batch normalization processing. And the ''same'' mode is used for padding in the convolution process.

In order to further reduce the amount of calculation, the improved method of feed-forward closed loop in Literature [25] is used to replace the two-layer convolution whose convolution kernel size is 3 with three-layer convolution whose convolution kernel size is 1, 3 and 1 respectively. In addition, the approximation precision is ensured by increasing the number of input and output channels. Take conv2_2 and conv2_3 as examples, and the replacement method is shown in deep model.

When the output channel of the convolution layer in the residual neural network changes, for example, the number of output channels from conv2_4 to conv3_1 changes from 64 to 128, the input and output of the feed-forward channel are inconsistent. And as shown in Figure 5, the input channel of T1 is 64, the output channel is 128. Therefore, the feed-forward channel and the tail convolution output cannot be added and activated directly, so the output channel needs to be expanded during the convolution operation of feed-forward. Take layer T1 as an example, the replacement of feed-forward closed loop and channel expansion structure is analyzed. And its similar structural replacements also include T2 and T3.

Compared with the two kinds of replaced structures, the number of input and output channels of the feed-forward convolution operation is 256. The input channel of the feed-forward convolution operation is 256, but the output channel is 512. Only after the channel is expanded, the output channel is the same as that of the Conv3_1_3.

C. DECODER STRUCTURE CONSTRUCTED BY DECONVOLUTION

The feature image is extracted from the angiographic image by deep residual neural network, which realizes the function of encoder. A convolution kernel of 2×2 is used on the image of 3×3, and the convolution process of stride $S = 1$ is easily dscribed, which will not analyze in detai [26].

The input image is vector $X = x_1, x_2, \cdots, x_9^T$ and the output feature image is $Y = y_1, y_2, y_3, y_4^T$. The convolution process can be expressed as follows:

$$CX = Y \tag{1}$$

$$C = \begin{pmatrix} \omega_1 & \omega_2 & 0 & \omega_3 & \omega_4 & 0 & 0 & 0 & 0 \\ 0 & \omega_1 & \omega_2 & 0 & \omega_3 & \omega_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_1 & \omega_2 & 0 & \omega_3 & \omega_4 & 0 \\ 0 & 0 & 0 & 0 & \omega_1 & \omega_2 & 0 & \omega_3 & \omega_4 \end{pmatrix} \tag{2}$$

In deep neural network, the process of deconvolution is the inverse of convolution. Therefore, the deconvolution of Equation (1) can be regarded as the propagation process from Y to X. Set the output loss function as Q, and calculate the back propagation of convolution according to BP (back propagation) algorithm [27], which is shown as follows:

$$\frac{\partial Q}{\partial x} = \left(\frac{\partial Q}{\partial x_1} \cdots \frac{\partial Q}{\partial x_9}\right)^T \tag{3}$$

According to Eq. (1), we can obtain:

$$\begin{aligned} \frac{\partial Q}{\partial x_i} &= \sum_{j=1}^{4} \frac{\partial Q}{\partial y_j} \cdots \frac{\partial y_j}{\partial x_i} \\ &= C_{1i}\frac{\partial Q}{\partial y_1} + C_{2i}\frac{\partial Q}{\partial y_2} + C_{3i}\frac{\partial Q}{\partial y_3} + C_{4i}\frac{\partial Q}{\partial y_4} \\ &= C_{all,i}^T \frac{\partial Q}{\partial y} \end{aligned} \tag{4}$$

where C_{ji} represents the element of row J and first column in the matrix; $C_{alli}^T = C_{1i} \cdots C_{4j}$. Therefore, the equation can be rewritten as follows

$$\frac{\partial Q}{\partial Y} = \left(\frac{\partial Q}{\partial y_1} \cdots \frac{\partial Q}{\partial y_4}\right)^T \frac{\partial Q}{\partial X} = C^T \frac{\partial Q}{\partial Y} \tag{5}$$

According to Eq. (4), the essence of deconvolution is to multiply the input by C^T, so deconvolution is also called transposed convolution.

Deconvolution is used to realize decoder function [27]. Decoder not only needs to segment the blood vessels through the extracted features, but also needs to restore the segmentation results to the original size of the input image. Therefore, the information source of deconvolution can't be limited to the feature output of encoder. On the basis of ensuring the accuracy of segmentation, in order to reduce the amount of calculation, the information source of decoder is determined as the output feature image of en-coder and the output result of conv4_4 after a large number of experiments. [29] And the two deconvolutions are fused to realize the segmentation of blood vessels. The implementation scheme of deconvolution to realize decoder function is shown in Fig. 6.

Firstly, deconvolution is performed on the output feature image of encoder, and the output channel is set as 2 (corresponding to 2 segmentation types); the output

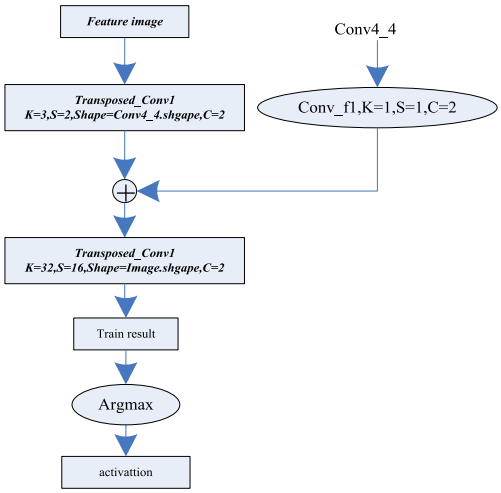


FIGURE 6. Decnnvnlutinn for decoder in deep model.

size corresponds to the output feature size of Conv4_4 convolution layer. Before the fusion of Conv4_4 feature output, a convolution operation with convolution kernel k = 1, s = 1, C = 2 is introduced to change the channels of the feature image into 2. After fusing the two feature information, deconvolution is performed again. And the output channel is set as 2, the output size is the original size of the angiographic image. At this time, the output specification is $[g_h g_w 2]$, that is, two matrices with the same size as the input image. If Softmax operation is introduced on this basis, the two matrices respectively represent the probability that each pixel belongs to the blood vessels and the background. When the training is over, the two channels of the output results are processed by Argmax operation to get the segmentation results of the blood vessels.

D. BATCH NORMALIZATION

The process of deep neural network training is to make the weight converge to the optimal value through error back propagation and random gradient descent method. With the increasing number of neural network layers, gradient dispersion or explosion, over-fitting and weight oscillation may be caused, which makes the training of neural network more difficult. The residual neural network used in Literature [26] can not only approximate the real model more accurately, but also restrain the gradient dispersion or explosion to a certain extent. To solve the over-fitting problem of training model, technology Dropout is usually used. In each training process, a certain proportion of weights are randomly selected not to participate in the training, so as to reduce the over-fitting. To deal with the problem of weight oscillation, the step length in the gradient descent process is gradually attenuated by using the weight decay technology, so as to approximate the optimal value of the weight accurately.

In the process of training neural network, the data flow is transferred layer by layer, and the change of the weight of the low-level network will inevitably cause the change of the distribution of the output data, which is one of the reasons for the difficulty of deep neural network training. Before inputting each convolution operation, the batch normalization technology is introduced to normalize the data, replacing dropout, weight decay and other technologies, which reduces the difficulty of neural network training.

Set all samples as $X = \{x_1, x_2, \cdots, x_N\}$; if all samples are normalized before per training, there is:

$$\hat{x}_i = \frac{x_i - E(X)}{\sqrt{Var(x_i)}} \tag{6}$$

where, $E(\cdot)$ is the mean value, and $Var(\cdot)$ is the variance. In each convolution operation, all training samples obey the same distribution. However, for the huge training samples of deep neural network, it is a very large calculation to calculate the mean and variance of all samples. Therefore, batch normalization is used instead of global mean and variance. That is, calculate the mean and variance of each training batch samples, and then take the mean of all mean and variance, and replace the global mean and variance. Set the number of training samples per time as m, and the implementation process of batch normalization is done, where only several lines of code are needed in TensorFlow framework.

V. EXPERIMENTS AND RESULTS ANALYSIS

A. DATA RESOURCES AND EXPERIMENTAL PLATFORM

In this paper, we use the coronary angiography image database as the object to carry out verification experiments. The angiographic image database has been accurately labeled for blood vessels and can be used as a training sample. Each sample pixel is 512 × 512. Due to the limitation of the GPU memory unit, each sample is cropped to a size of 64×64 pixels, and there are 4 500 samples in total, where 4 480 are used as training samples and 20 are used as test samples [23].

The experimental platform is equipped with Intel-i7-7700K quad-core CPU processor, 32G memory, ASUS STRIX-GTX1080TI-11G graphics card, deep learning framework using Google's Tensor Flow.

B. DATA PREPROCESSING AND EVALUATION INDICATORS

Sample data is pre-processed before training the neural network. In order to quantitatively evaluate the segmentation results, Recall Rate, Precision Rate, and F-measure are adopted as evaluation indicators to analyze the segmentation results. Their equations are written as follows:

$$\mathrm{Re}call = \frac{B_{seg}}{B_{seg} + I_{unseg}} \tag{7}$$

$$\mathrm{Pr}ecision = \frac{B_{seg}}{B_{seg} + I_{wseg}} \tag{8}$$

$$F\_measure = \frac{2\mathrm{Re}call \times \mathrm{Pr}ecision}{\mathrm{Re}call + \mathrm{Pr}ecision} \tag{9}$$

where B_{seg} is denoted as the number of pixels with the correct blood vessel in the segmentation result, I_{unseg} is denoted the number of pixels in the image that are blood vessels but not recognized as blood vessels, I_{wseg} indicates the number of

pixels in the image that incorrectly recognize the background as blood vessels.

The recall rate represents the ratio of the pixel points that are segmented into blood vessels to the true blood vessel pixels, characterizing the accuracy of the blood vessel segmentation without considering the background of the coronary angiography image. The accuracy rate represents the ratio of pixels that are correctly segmented into blood vessels to all pixels that are segmented into blood vessels. A high accuracy rate means that the blood vessels can be fully extracted. The F value is an evaluation index of the two indicators of comprehensive recall rate and accuracy rate, which is used to reflect the overall index.

According to the existing deep learning frameworks, there are two typical network structures: VGG-based full convolution neural network and deep network based on VGG and fully connected conditional random field. In order to verify the effectiveness of our proposed model in this paper, a comparative experiment with VGG full convolution neural network (VGG) and VGG conditional random field network (VGGCRF) was carried out on the iaild database. For the sake of simplicity, our proposed model in this paper is written by RESNET. In the experiment, VGG adopts the structure shown in literature [33], and the convolution kernel parameters in the first 13 layers of neural network uses the trained values. The convolution kernels of F6, F7 and F8 in the full connection layer are set to [16, 16, 512, 1024], [1, 1, 1024, 2048], [1, 1, 2048, 2], respectively. Finally, the multi-layer deconvolution results are fused to achieve the output of blood vessel segmentation results. VGGCRF is the fully connected conditional random field in the last layer of the VGG model. The specific structure is described in the literature [25].

C. COMPARATIVE EXPERIMENT AND RESULT ANALYSIS

The training process of deep neural network is to make the weight of network converge by learning the sample data. The three network structures all use the cross-entropy as the loss function, and the training process is shown in Figure 7. Because of the special structure of residual neural network and batch normalization technology, the weights of neural network are easier to train and the convergence performance of weights is better.

It should be noted that VGGCRF training is still difficult even when batch normalization technology is used. We think that VGGCRF can be trained in two steps to make the weight of the network converge. First of all, VGG network training is used directly to realize the rough segmentation of blood vessels, and then the full connection condition is introduced to participate in the training until the training stop condition is met. The training time of the three networks is shown in Table 1.

Because RESNET network structure is more complex than VGG and VGGCRF, the network training time is the longest, but the training time of a single sample is the shortest. In VGGCRF, full connection conditional random fields

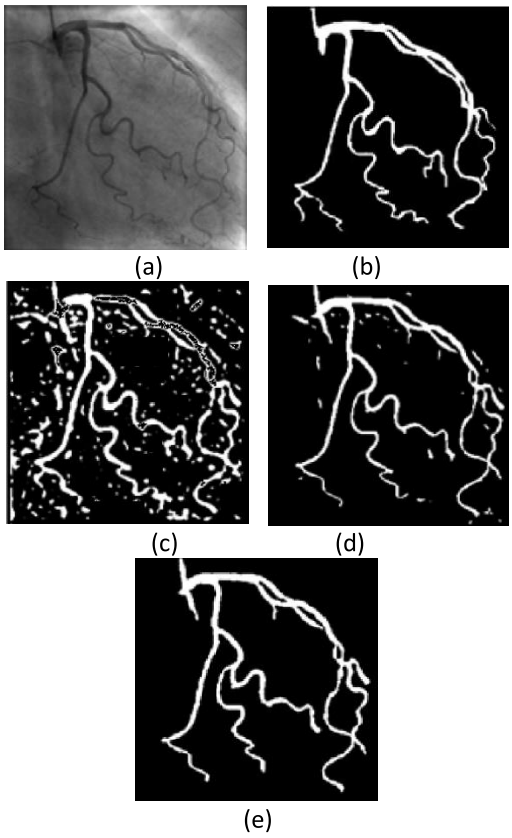


FIGURE 7. Qualitative results for different model. (a) Raw coronary arteriography image; (b) Benchmark image; (c) VGG model; (d) VGGCRF model (e); Proposed model.

TABLE 1. Time for training the deep neural networks.

Time	VGG	VGGCRF	Proposed
Calling time for network	32.21	35.29	143.52
Single sample	0.99	5.36	0.43
Training time for sample-set	3328	23525	1452

are introduced. It is necessary to iteratively calculate the pair potential function between all pixel-pairs, so the training time is the longest.

The experiment selected three kinds of typical angiographic blood vessels including complex blood vessels with low contrast, complex blood vessels with fuzzy image and regular distributed dim-small blood vessels. The above-mentioned deep neural network models are used for segmentation experiments, and the experimental results were compared and analyzed, where the angiography image shown in Fig. 8 contains a complicated background.

According to the segmentation results, all three networks can segment blood vessels, and VGG can roughly determine blood vessels within a certain range, but the accuracy of blood vessel edge information is not high. VGGCRF has improved the extraction of blood vessel edges compared

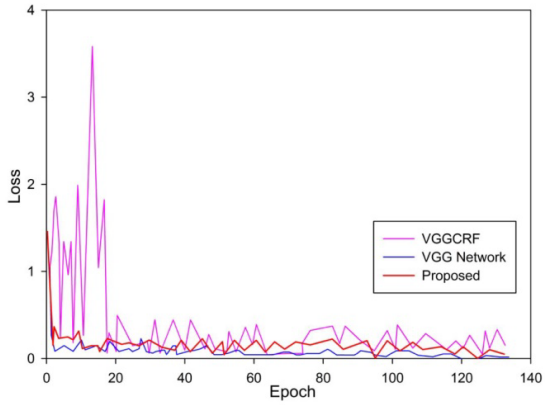


FIGURE 8. Convergence performance for different models.

TABLE 2. Quantitative result.

Model	VGG model	VGGCRF model	Proposed model
Recall	0.504	0.639	0.868
Precision	0.749	0.694	0.836
F-measure	0.603	0.665	0.851

to VGG. ResNet can better extract the block information and edge features of blood vessels. At the same time, the three networks have a certain degree of mis-segmentation, where VGG is dominated by image contour information, and the misclassification result is dense information; VGGCRF mitigates the occurrence of misclassification on the basis of VGG; ResNet can process the patch information of blood vessels. As shown in Figure 9, RESNET can significantly improve the recall rate, accuracy rate and F value of vascular segmentation in complex angiographic image.

In Figure 10, the distribution of blood vessels is relatively regular, and the interference of blood vessel segmentation mainly comes from the influence of background. RESNET network structure is more accurate to detect the contour of blood vessels, and it can segment the smaller vessels effectively [30]. According to the segmentation result evaluation index shown in Figure 10, the recall rate, accuracy rate and F value of RESNET have been greatly improved.

Figure 10 is the angiographic image of a single vessel with complex structure and large area of shadow. In addition, the gray level of some blood vessels is similar to its background, so the influence of background is more obvious. It can be seen from the segmentation results that there are large areas of mis-segmentation in the three network structure segmentation results. However, RESNET successfully avoids the influence of uneven background, and the edge detection of blood vessels is more accurate [30]. However, compared with VGG, RESNET is not robust to shadow interference. It mistakenly recognize the large staggered shadow as background, resulting in low segmentation accuracy and low F value.

Comparing the segmentation results of three different networks, it can be seen that VGG can roughly recognize the

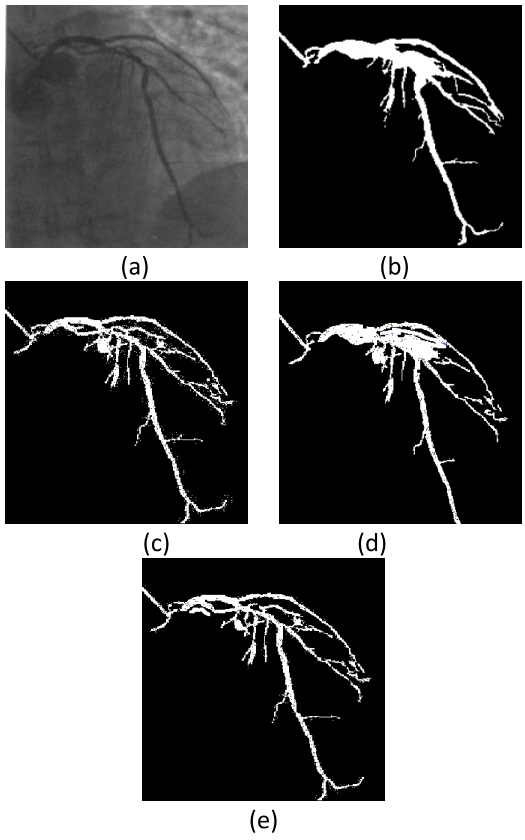


FIGURE 9. Qualitative results for different model.(a)Raw coronary arteriography image; (b) Benchmark image; (c) VGG model; (d) VGGCRF model (e); Proposed model.

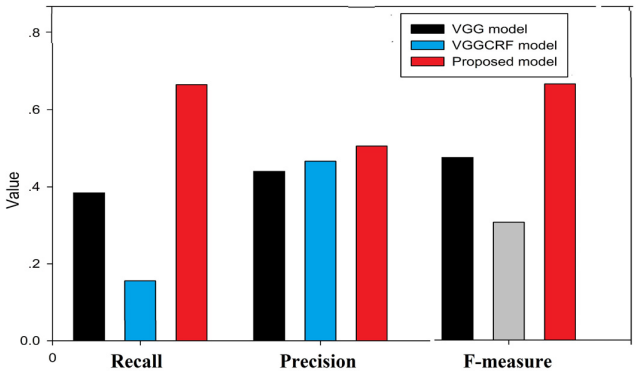


FIGURE 10. Performance index of building segmentation.

range of blood vessels, and also can better extract the edge of vessels with less interference; VGGCRF is very accurate to extract the edge of vessels with straight line and no interference, but there is obvious missegmentation for other types of edges. RESNET is more accurate in the edge extraction of blood vessels, and the segmentation results are patch shape. Compared with VGG and VGGCRF, RESNET is more accurate in the segmentation of blood vessels.

VI. CONCLUSION

Aiming at the problem of automatic and accurate segmentation of blood vessels in angiographic images, this paper

proposes a framework of encoder-decoder feature extraction and vessel segmentation based on deep residual neural network, and carries out experimental verification on angiographic database by means of batch normalization technology. The experimental results show that the proposed algorithm in this paper has high efficiency, and the training time for a single sample is 0.23 s, and the training time for a sample set is 1418.862 s. In the experimental results of complex blood vessels, regular blood vessels and single complex blood vessels, the segmentation accuracy is 0.837, 0.892 and 0.630 respectively; the F values are 0.851, 0.879 and 0.730 respectively. However, for complex vessels with low contrast and large area shadow, the proposed algorithm still has the problems of edge detection error and low segmentation accuracy. In the future work, we will focus on how to eliminate the interference in complex blood vessels and further improve the accuracy of blood vessel segmentation.

REFERENCES

[1] S. Takashio, M. Yamamuro, Y. Izumiya, S. Sugiyama, S. Kojima, E. Yamamoto, K. Tsujita, T. Tanaka, S. Tayama, K. Kaikita, S. Hokimoto, and H. Ogawa, "Coronary microvascular dysfunction and diastolic load correlate with cardiac troponin t release measured by a highly sensitive assay in patients with nonischemic heart failure," *J. Amer. College Cardiol.*, vol. 62, no. 7, pp. 632–640, 2013.

[2] F. C. Ciftci, M. Caliskan, O. Ciftci, H. Gullu, A. Uckuyu, E. Toprak, and F. Yanik, "Impaired coronary microvascular function and increased intima–media thickness in preeclampsia," *J. Amer. Soc. Hypertension*, vol. 8, no. 11, pp. 820–826, 2014.

[3] G. Tsujimoto, "Impaired coronary microvascular function in diabetics," *Ann. Nucl. Med.*, vol. 14, no. 3, pp. 165–172, 2000.

[4] G. Niccoli, G. Scalone, A. Lerman, and F. Crea, "Coronary microvascular obstruction in acute myocardial infarction," *Eur. Heart J.*, vol. 37, no. 13, pp. 1024–1033, 2016.

[5] K. Kakuta, K. Dohi, Y. Sato, T. Yamanaka, M. Kawamura, S. Nakamori, R. Okamoto, E. Fujii, N. Yamada, and M. Ito, "Coronary microvascular dysfunction and coronary artery calcification in patients with systemic lupus erythematosus, systemic sclerosis, and rheumatoid arthritis," *J. Amer. College Cardiol*, vol. 65, no. 10, p. A1674, 2015.

[6] S. S. Dhawan, R. P. A. Nanjundappa, P. Eshtehardi, M. Corban, L. Golub, L. Timmins, M. McDaniel, A. Quyyumi, and H. Samady, "Impaired coronary microvascular function is associated with features of plaque vulnerability," *J. Amer. College Cardiol.*, vol. 59, no. 13, p. E408, 2012.

[7] P. M. Elliott, H. Kindler, J. S. Shah, B. Sachdev, O. E. Rimoldi, R. Thaman, and P. G. Camici, "Coronary microvascular dysfunction in male patients with Anderson–Fabry disease and the effect of treatment with a galactosidase A," *Heart*, vol. 92, no. 3, pp. 357–360, 2006.

[8] H. Nakashima, Y. Akiyama, H. Tasaki, Y. Honda, T. Katayama, and K. Yano, "Coronary microvascular dysfunction in coronary artery disease associated with glucose intolerance," *J. Cardiol.*, vol. 30, no. 2, pp. 59–65, 1997.

[9] G. J. Ughi, T. Adriaenssens, K. Onsea, P. Kayaert, C. Dubois, P. Sinnaeve, M. Coosemans, W. Desmet, and J. D'hooge, "Automatic segmentation of in-vivo intra-coronary optical coherence tomography images to assess stent strut apposition and coverage," *Int. J. Cardiovascular Imag.*, vol. 28, no. 2, pp. 229–241, Feb. 2012.

[10] F. Lugauer, J. Zhang, Y. Zheng, J. Hornegger, and B. M. Kelm, "Improving accuracy in coronary lumen segmentation via explicit calcium exclusion, learning-based ray detection and surface optimization," *Proc. SPIE*, vol. 9034, pp. 22–38, Mar. 2014.

[11] Y. Wang and P. Liatsis, "An automated method for segmentation of coronary arteries in coronary CT imaging," in *Proc. Develop. E-Syst. Eng.*, Sep. 2010, pp. 2652–2668.

[12] D. Han, H. Shim, B. Jeon, Y. Jang, Y. Hong, S. Jung, S. Ha, and H.-J. Chang, "Automatic coronary artery segmentation using active search for branches and seemingly disconnected vessel segments from coronary CT angiography," *PLoS ONE*, vol. 11, no. 8, 2016, Art. no. e0156837.

[13] M. M. Jawaid, B. S. Chowdhry, and G. Slabaugh, "Automated framework for CTA coronary segmentation and quantitative validation," in *Proc. Int. Conf. Innov. Elect. Eng. Comput. Technol. (ICIEECT)*, Apr. 2017, vol. 24, no. 15, pp. 6–18.

[14] G. Santini, D. D. Latta, N. Martini, G. Valvano, A. Gori, A. Ripoli, C. L. Susini, L. Landini, and D. Chiappino "An automatic deep learning approach for coronary artery calcium segmentation," in *Proc. Eur. Med. Biol. Eng. Conf.*, 2018, vol. 24, no. 15, pp. 52–68.

[15] M. Pazdernik, Z. Chen, H. Bedanova, J. Kautzner, V. Melenovsky, V. Karmazin, I. Malek, A. Tomasek, E. Ozabalova, J. Krejci, J. Franekova, A. Wahle, H. Zhang, T. Kovarnik, and M. Sonka, "Detecting early cardiac allograft vasculopathy using highly automated 3D coronary optical coherence tomography segmentation analysis," *J. Heart Lung Transplantation*, vol. 37, no. 4, pp. S105–S106, 2018.

[16] M. Schaap, T. van Walsum, L. Neefjes, C. Metz, E. Capuano, M. de Bruijne, and W. Niessen, "Robust shape regression for supervised vessel segmentation and its application to coronary segmentation in CTA," *IEEE Trans. Med. Imag.*, vol. 30, no. 11, pp. 1974–1986, Nov. 2011.

[17] U. Sivalingam, M. Wels, M. Rempfler, S. Grosskopf, M. Suehling, and B. H. Menze, "Inner and outer coronary vessel wall segmentation from CCTA using an active contour model with machine learning-based 3D voxel context-aware image force," *Proc. SPIE*, vol. 9785, pp. 99–108, Mar. 2016.

[18] Y. L. Yong, L. K. Tan, R. A. Mclaughlin, K. H. Chee, and Y. M. Liew, "Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography," *J. Biomed. Opt.*, vol. 22, no. 12, pp. 1052–1065, 2017.

[19] M. Breeuwer, P. Ermes, and B. Gerber, "Clinical evaluation of automatic whole-heart and coronary-artery segmentation," *J. Cardiovascular Magn. Reson.*, 2009, vol. 11, no. 1, pp. 1–12.

[20] D. Lesage, E. D. Angelini, G. Funka-Lea, and I. Bloch, "Adaptive particle filtering for coronary artery segmentation from 3D CT angiograms," *Comput. Vis. Image Understand.*, vol. 151, no. 8, pp. 29–46, 2016.

[21] C. Feng and Y. Hu, "Segmentation of coronary artery using region based level set with edge preservation," *J. Med. Imag. Health Informat.*, vol. 6, no. 7, pp. 1727–1731, 2016.

[22] Z. Liu, X. Li, P. Luo, C. C. Loy, and X. Tang, "Deep learning Markov random field for semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1814–1828, Aug. 2018.

[23] C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu, "An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation," in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges* (Lecture Notes in Computer Science), vol. 10663, M. Pop *et al.*, Eds. Cham, Switzerland: Springer, 2018.

[24] G. Santini, D. D. Latta, N. Martini, G. Valvano, A. Gori, A. Ripoli, C. L. Susini, L. Landini, and D. Chiappino, "An automatic deep learning approach for coronary artery calcium segmentation," in *Proc. Nordic-Baltic Conf. Biomed. Eng. Med. Phys.*, 2018, pp. 374–377.

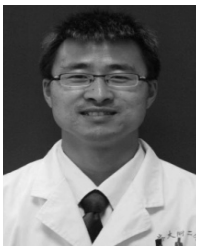
[25] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia-Rodriguez, "A survey on deep learning techniques for image and video semantic segmentation," *Appl. Soft Comput.*, vol. 70, no. 61, pp. 253–266, 2018.

[26] A. King, S. M. Bhandarkar, and B. M. Hopkinson, "A comparison of deep learning methods for semantic segmentation of coral reef survey images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 123–133.

[27] M. Wang, L.-L. S. Ong, J. Dauwels, and H. H. Asada, "Multicell migration tracking within angiogenic networks by deep learning-based segmentation and augmented Bayesian filtering," *J. Med. Imag.*, vol. 12, no. 5, pp. 25–33, 2018.

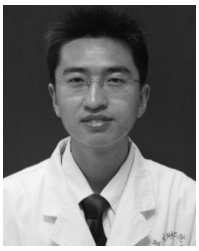
[28] M. Li, Q. Yin, and M. Lu, "Retinal blood vessel segmentation based on multi-scale deep learning," in *Proc. Federated Conf. Comput. Sci. Inf. Syst.*, Sep. 2018, pp. 698–708.

[29] M.-X. Li, S.-Q. Yu, W. Zhang, H. Zhou, X. Xu, T.-W. Qian, and Y.-J. Wan, "Segmentation of retinal fluid based on deep learning: Application of three-dimensional fully convolutional neural networks in optical coherence tomography images," *Int. J. Ophthalmol.*, vol. 12, no. 6, pp. 1012–1020, 2019.



SHIWEN PAN received the Master of medicine degree in imaging medicine and nuclear medicine with The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.

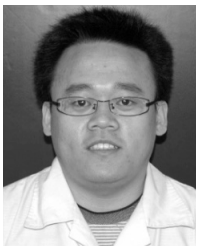
GUOHUA FAN received the Doctor of medicine degree in clinical medicine from The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.



WEI ZHANG received the Master of medicine degree in imaging medicine and nuclear medicine from The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.

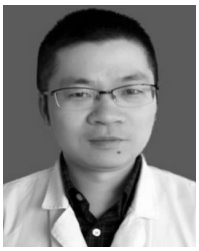
JIANPING GONG received the Doctor of medicine degree in majoring in clinical medicine from The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.

WANJUN ZHANG received the Master of medicine in imaging medicine and nuclear medicine degree from The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.



BO ZHANG received the Master of medicine degree in imaging medicine and nuclear medicine from The Second Affiliated Hospital of Soochow University, China. His research interest includes imaging diagnosis.

LIANG XU received the Doctor of medicine degree in clinical medicine, from the Second Affiliated Hospital of Soochow University, in China. His research interest includes imaging diagnosis.



HAIBO GU received the Master of medicine in cardiology degree from The Second Affiliated Hospital of Soochow University, China. His research interest includes cardiovascular disease.