IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 11, 2018, accepted November 27, 2018, date of publication December 17, 2018,
date of current version January 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2886009

Segmentation of the Main Vessel of the Left Anterior Descending Artery Using Selective Feature Mapping in Coronary Angiography

KYUNGMIN JO[1], JIHOON KWEON[1], (Member, IEEE), YOUNG-HAK KIM[2], AND JAESOON CHOI[1,3], (Member, IEEE)

[1]Biomedical Engineering Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul 05505, South Korea
[2]Department of Cardiology, University of Ulsan College of Medicine, Asan Medical Center, Seoul 05505, South Korea
[3]Department of Biomedical Engineering, University of Ulsan College of Medicine, Seoul 05505, South Korea

Corresponding author: Jaesoon Choi (fides@amc.seoul.kr)

This work was supported in part by the National Research Foundation of Korea funded by the South Korean Government (MSIT) under Grant NRF2017R1A2B3009800, and in part by the Korea Health Technology Research and Development Project through the Korea Health Industry Development Institute funded by the Ministry of Health and Welfare, South Korea, under Grant HI17C2410.

ABSTRACT X-ray angiography, used in the evaluation of coronary artery disease, presents difficulties in the performance of quantitative coronary angiography analysis, by identifying major vessels. These difficulties are due to problems such as non-uniform illumination, low contrast ratio, and the presence of other tissues. Therefore, segmentation of the desired vessels in images containing multiple blood vessels is clinically important. This paper proposes selective feature mapping as a method for segmenting the left anterior descending artery main vessel in coronary angiography images. The proposed method consists of two steps for generating a candidate area of an image and then segmenting it. To generate the candidate area, feature maps that overlap significantly with the area of the ground truth are selected and combined. Segmentation then is performed using a neural network that learns only the ground truth region of the input image. The proposed method consists of eight modules: pre-processing of the angiogram, resizing of ground truth, pre-processing for segmentation, post-processing for segmentation, network, and segmentation network. This method has a precision of about 0.066, recall of 0.091, and an F1 score of 0.094, values which are higher than those generated by the U-Net, one of the conventional techniques.

INDEX TERMS Angiogram, convolutional neural networks, deep learning, left anterior descending artery, selective feature map, selective filter, vessel segmentation.

I. INTRODUCTION

Coronary artery disease (CAD) is one of the world's leading causes of death, and its incidence is rapidly increasing worldwide [1]–[3]. This disease is caused by the narrowing of blood vessels due to blockage by a plaque consisting of fat, cholesterol, and calcium [1], [2]. X-ray angiography is the primary imaging method for evaluating CAD [1], [2], [4]. A cardiologist looks at an image, identifies the stenosis, and determines the severity of the disease [1]–[3]. However, X-ray angiography has a large inter-individual variation in its analytical results due to problems such as non-uniform illumination, low contrast ratios, low signal-to-noise ratios, the presence of other tissues, and camera motion [2], [3]. To perform quantitative coronary angiography (QCA) analysis with high accuracy, considerable experience is required [1]–[3]. In particular, in clinical studies, QCA requires the identification and analysis of major blood vessels, and thus relies on the manual labor of experts. It is therefore important to consistently detect affected blood vessels among the multiple blood vessels in an image. The purpose of this study was to automatically segment the main vessels of the left anterior descending (LAD) artery, which generally supplies the widest area of myocardium from coronary angiography images (Fig. 1).

Considerable research has been conducted into the segmentation of blood vessels automatically, or semi-automatically, from an angiography image using traditional methods. In one study [1] blood vessels were segmented from images in three stages: Hessian filter enhancement; feature extraction; and vessel region detection. Another study [2]

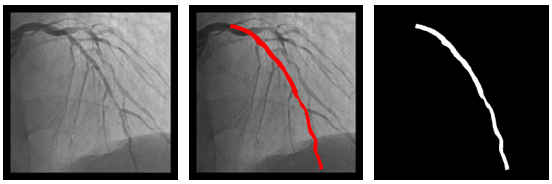

FIGURE 1. LAD main vessel segmentation among vessels in an angiography image (a: angiography image, b: overlapped ground truth over LAD main vessel of angiography image, c: ground truth).

used three different super-pixel scales to segment blood vessels by determining the probability of blood vessels for each super-pixel. Another study determined vessel structure using the multiscale probability of blood vessels in each super-pixel. In addition, these authors used Gabor filters to segment coronary arteries by binarizing the Gabor filter results using a new thresholding technique [4].

Deep learning techniques have been used to detect or segment blood vessels in angiography images. The research reported in [3] improved contrast via pre-processing, and then used about 1,040,000 patches of pixels to distinguish the vessel of interest from the background. In addition, noisy low-level binary segmentation and optical flow based on U-Net has been used to generate multiclass annotations, and multistage segmentation was then used to distinguish blood vessels [5]. There is a study reporting research into improving the image quality of an input image by using a patch around each pixel as input to a trained convolutional neural network (CNN) to distinguish whether each pixel is part of a blood vessel or is background [6]. These studies show fairly high performance in vessel recognition, but aim to segment all blood vessels from the image.

Studies on extracting blood vessels from retinal images and classifying them as arteries or veins have also been reported. The research reported in [7] detected blood vessels using vessel profile-based feature vectors and vessel ROI-based feature vectors, and performed classification using support vector machines and neural networks. Classification has also been performed using a convolutional neural network architecture with three convolutional layers and three fully-connected layers [8]. These techniques also produce data for all vessels. With these approaches, it is important to use accurate datasets, because their performance depends largely on the accuracy of the dataset. However, it is time-consuming and burdensome to generate ground truth data for all vessels in angiography images.

In this paper, we propose selective feature mapping (SFM) as a new method for segmenting the main blood vessels of LAD in a coronary angiography image. As an extension of the CNN, it selects several feature maps that best match a label from among the feature maps produced by the CNN's last layer. We apply convolution, average pooling, and thresholding to the last layer of the CNN to select some of the filters that make up the last layer, and combine only the results from these filters to eliminate unnecessary feature maps. Unlike the previous studies mentioned above, the proposed method reduces the effort required for generating datasets because it only requires the ground truth of the main vessel of the LAD.

The composition of this paper is as follows. Section 2 describes the proposed method which consists of two stages: (1) SFM to find a candidate areas and (2) segmentation of the main LAD vessel. Stage two is made up of eight modules: pre-processing of the angiogram; resizing ground truth; preprocessing for segmentation; post-processing for segmentation; network; segmentation of the network; and filtering. Each is explained in turn. Section 3 presents the experimental conditions, performance criteria and results, and Section 4 discusses our conclusions and outlines future work.

II. METHODS

This section describes the proposed method in more detail. It consists of two steps: (1) selective feature mapping to find a candidate area; and (2) segmenting the main vessel from the candidate area. Figures 2 and 3 summarize the training and test structure of the proposed method.

The main aim of the first step, selective feature mapping, is to automatically select filters that detect the main vessel in the input image (the FS module), and then combine the corresponding feature maps to generate the candidate area (the FM module). In order to train filter selection automatically according to the characteristics of the image, we generate an answer filter from a ground truth image (the FS_GT module) as shown in Fig. 2 (a).

The second step is to segment the main vessel from the candidate area. Specifically, the main vessel is segmented from the candidate area generated by the selective feature mapping (Fig. 3 (b)). During training, we set the candidate area within a certain range from the ground truth area and segment the main vessel from this area through the segmentation network (Fig. 2 (b)). In other words, the input image for training and testing is different in the second stage, so pre-processing and post-processing for segmentation are included in the testing stage to compensate for this difference.

There are eight modules that make up the first and second steps (the blue boxes in Fig. 2 and 3): pre-processing of the angiogram; resizing of the ground truth; FS; FM; pre-processing for segmentation; post-processing for segmentation; the network; and the segmentation network.

This section describes these modules individually, except for the network and segmentation network, and finally describes the three losses (the orange boxes in Fig. 2).

The reason for not providing detailed descriptions of networks and segmentation networks is that they can use an existing network algorithm. We used U-Net as the segmentation network for performance comparison.

A. PRE-PROCESSING OF THE ANGIOGRAM

In Steps 1 and 2, the Frangi filter [9] is applied, to highlight the vascular region in an angiography image. It is based on the Hessian matrix. The Hessian matrix extracts

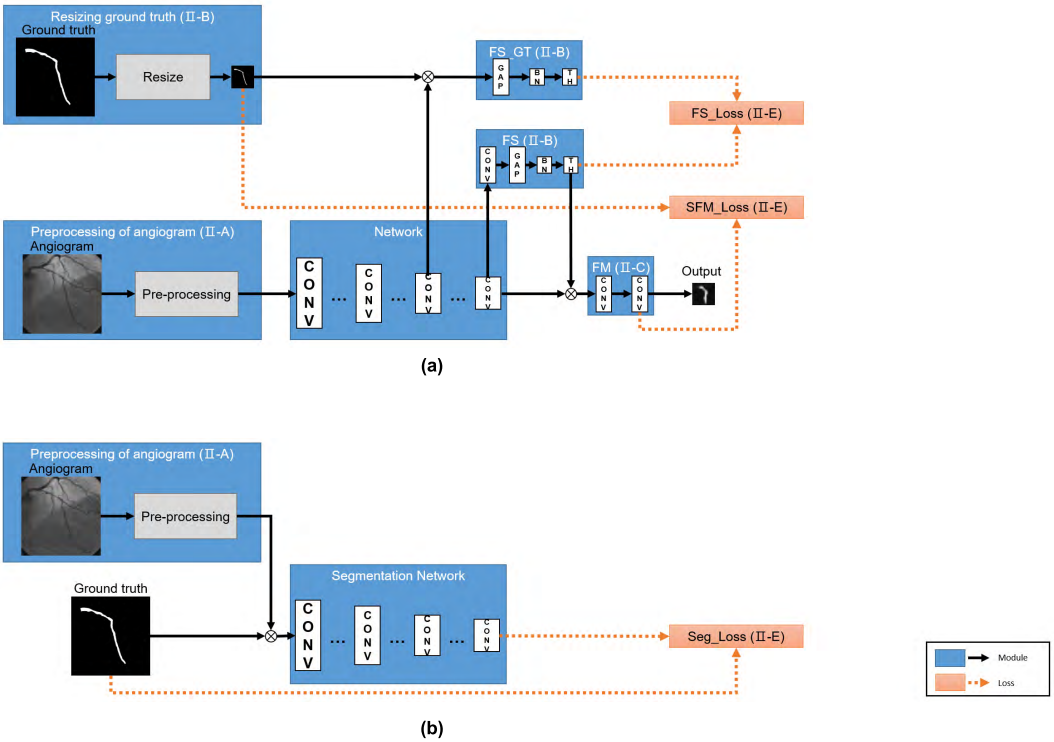

FIGURE 2. Summary of training architecture of the proposed method (blue boxes: modules, orange boxes: loss functions, gray boxes: processing, orange dotted arrow line: calculating loss function). (a) Selective Feature Mapping (Training) – Aim: candidate area generation. (b) Segmentation (Training) – Aim: LAD main vessel segmentation within candidate area.

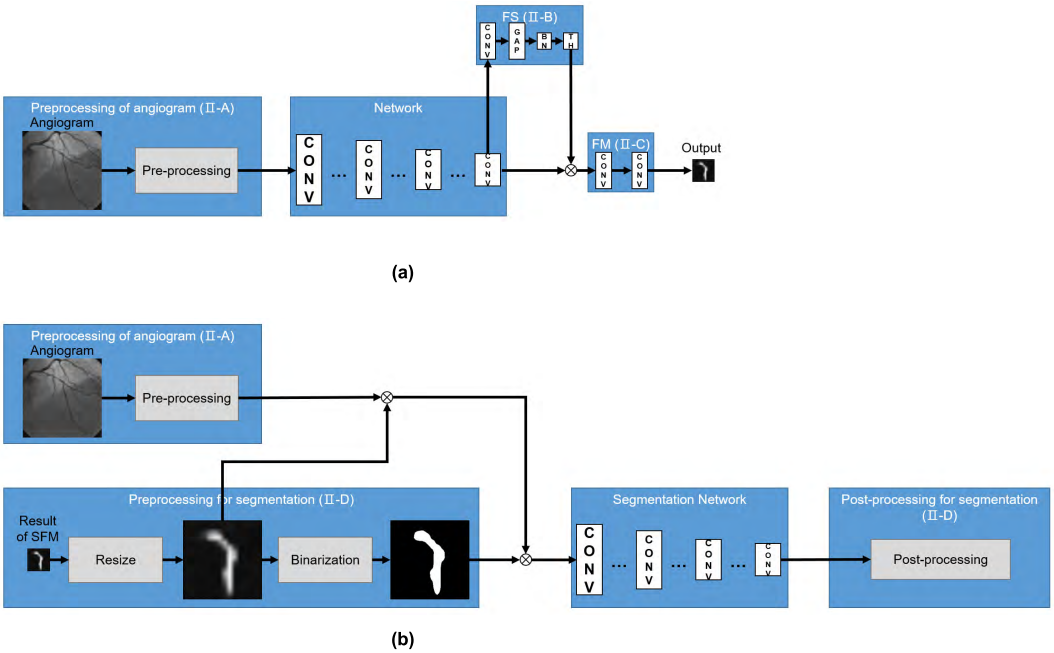

FIGURE 3. Summary of test architecture of the proposed method (blue boxes: modules, orange boxes: loss functions, gray boxes: processing, orange dotted arrow line: calculating loss function). (a) Selective Feature Mapping (Test) – Aim: candidate area generation. (b) Segmentation (Test) – Aim: LAD main vessel segmentation within candidate area.

eigenvectors and eigenvalues from each pixel of the input image via a second derivative. As shown in Fig. 4, the blood vessel and the background are classified using the eigenvalue (λ_2) of the largest change in direction in each pixel and the eigenvalue (λ_1) perpendicular thereto. The eigenvalue of a blood vessel has at least one non-zero component.

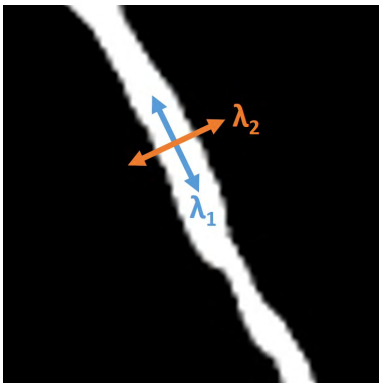

FIGURE 4. Eigenvectors in blood vessel image (Orange arrow: direction with the largest change in pixel value, Blue arrow: direction perpendicular to the orange arrow).

However, there is no specific directionality in the background, so eigenvalues in all directions of the background approach zero. Thus, the background can be removed from the image using eq. 1.

$$\mathrm{S} = \|\mathrm{H}\|_{\mathrm{F}} = \sqrt{\sum_{j \le D} \lambda_j^2} \tag{1}$$

S is the second order structureness, H is the Hessian matrix, and D is the image dimension. The definition of the Frangi filter of a two dimensional image is as follows.

$$\mathrm{f(x)} = \begin{cases} 0, & \text{if } \lambda_2 > 0 \\ \exp\left(-\dfrac{R_B^2}{2\beta^2}\right)\left(1 - \exp\left(\dfrac{S^2}{2c^2}\right)\right), & \text{otherwise} \end{cases} \tag{2}$$

$\mathrm{R}_B = |\lambda_1| / |\lambda_2|$ describes deviations from blob-like structures. β and c are thresholds that control the sensitivity of the line filter to the measured values R_B and S. Fig. 5 shows the results of the Frangi filter when β and c have a value between 0 and 1. As c increases, the less directional part gradually becomes the background. Conversely, the larger β is, the better the perception of thin sections of blood vessels and the better the perception of blood vessels.

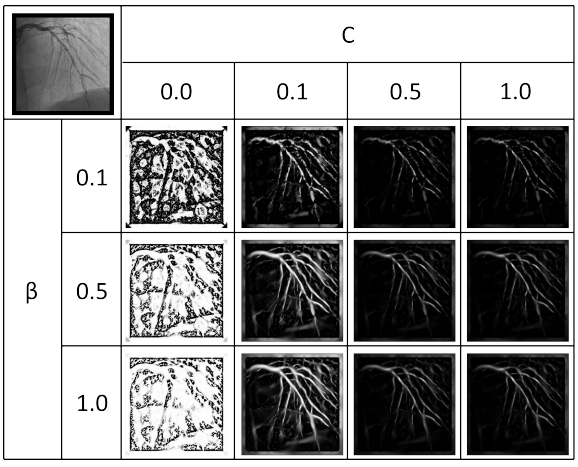

FIGURE 5. The result of Frangi filter with various values of β and c (left top: angiography image).

B. FILTER SELECTION AND RESIZING OF GROUND TRUTH

The general structure of a CNN consists of a convolutional layer and a pooling layer. In the convolutional layer, each filter detects a pattern of interest in the entire image area, through convolution. The feature map is an output image composed of the patterns detected from each filter. The pooling layer reduces the size of the input feature map and reduces the total number of parameters required for CNN training. The CNN repeatedly performs this process to detect a region of interest (ROI). Thus, from the feature map, we can observe the area of interest for each filter [10]. The heat map depicted in Fig. 6 shows several representative feature maps of the last convolutional layer after training on the network in Step 1. Using the heat map, the ROI of each feature map can be identified [11]. More specifically, in the heat map red represents the region of interest for each filter, and blue corresponds to the region of non-interest. Since each filter has different regions of interest, we can select feature maps where the ROI overlaps the main vessel of the LAD.

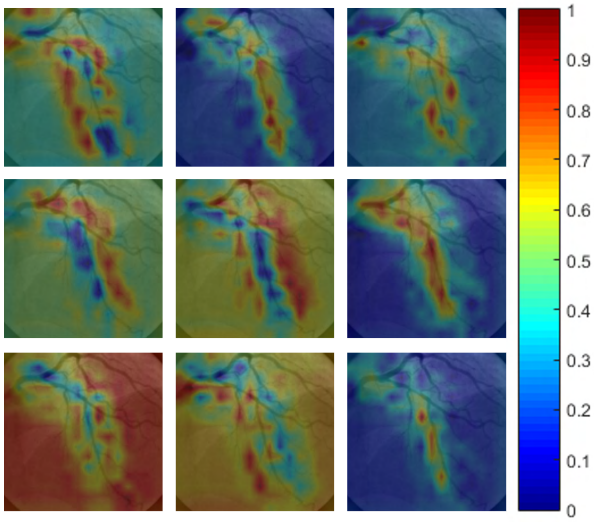

FIGURE 6. Heatmaps of several representative feature maps in last convolutional layer in the network (The closer to red, the area of interest for each filter, and the closer to blue, the less interesting area).

Filter selection involves choosing a filter that generates a feature map of the ROI from the last convolution layer of the network. This idea was inspired by filter pruning [20], which is performed by estimating the importance of a filter at each convolutional layer, to detect drowsiness. This method performs filter pruning using l1 normalization, fully connected layer, ReLU, and a sigmoid activation function. The proposed method performs filter selection through the convolutional layer, global average pooling, batch normalization, and thresholding (Fig. 2 and 3 (a)). Also, in our method,

a filter is selected that produces a feature map of interest in the ground truth area in the last convolution layer of the network. For this purpose, the size of ground truth is reduced to be the same as that of the feature maps of the last convolutional layer in the network. The feature maps of the last convolutional layer are then multiplied by the resized ground truth so that each feature map has only the value corresponding to the ground truth region. Feature maps are passed through the global average pooling (GAP) layer [12] so that each filter has one representative value. The range of the representative value is then adjusted using batch normalization [13], and a filter having a value equal to or higher than a certain threshold value is selected. This process is shown in FS_GT in Fig. 2 Through this process, the network can select a filter that involves the ground truth area. However, FS_GT cannot be used, since there is no ground truth in the test. To solve this problem, FS is added after the convolutional layer of the network, so that the filter can be selected automatically during the training and test procedures. FS is designed to have a structure parallel to that of FS_GT, and the result of FS_GT is regarded as the ground truth of FS. The input to FS_GT is the result of element-wise multiplication between the resized ground truth and the feature maps of the last convolutional layer. In order for FS to have the same structure, one convolutional layer is added before the GAP layer. The convolutional layer is always followed by an application of an activation function. The structure of the remaining FS is designed to be the same as FS_GT.

C. FEATURE MAPPING

In the selective feature mapping, feature maps corresponding to the selected filter in the FS are obtained from the network's last convolutional layer. Specifically, filters selected via the threshold layer, which is the last stage of the FS, have a value of 1, and the remaining filters have a value of 0. By multiplying this value by the feature maps of the last convolutional layer, only the feature maps corresponding to the selected filter survive. Thereafter, the number of feature maps is reduced through two convolutional layers with activation functions, and one image is generated.

D. PRE-PROCESSING AND POST-PROCESSING FOR SEGMENTATION

Fig. 2 (b) shows training using the ground truth to indicate the position of the LAD main vessel. However, since there is no ground truth in the test, we use the SFM result as an indicator. Nonetheless, the size of the result of SFM is equal to the size of the last convolutional layer, and is smaller than the angiogram. Therefore, resizing is important. However, this process results in pixels with values between 0 and 1, unlike the ground truth. Therefore, we apply Otsu's method to generate binary masks from the results of the resized SFM results. The generated binary mask has a wider white area than the ground truth. To compensate for this difference, we multiply the angiogram by the resized value. This makes the pixels with values close to 1 in the resized image more likely to be part of the main vessel of the LAD. Through these processes, the final mask plays a role corresponding to the ground truth in Fig. 2 (b).

Unlike training, post-processing is used in testing. Since training uses ground truth, there is no false detection area. However, in testing, the false detection area exists because the SFM result is used as ground truth. Therefore, in the post-processing for the segmentation module, a small area is removed from the output of the segmentation network. The size of the small area is determined based on the average value of the isolation area in the output image.

Fig. 7 shows the results corresponding to each step of Fig. 3 (b). As shown in Fig. 7, the peripheral blood vessels of Fig. 7 (c) were removed from the candidate images (Fig. 7 (d)).

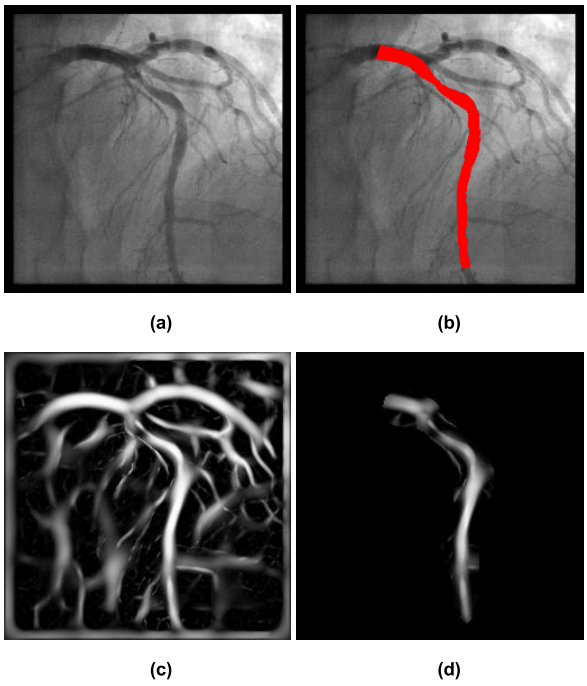

FIGURE 7. The results of each block in flowchart of preprocessing for segmentation (a: angiogram, b: overlapped ground truth over LAD main vessel of angiogram, c: the result of pre-processing of angiogram (Frangi parameter : β is 1.0, c is 0.1), d: filtered angiogram, which is the input of the segmentation network).

E. LOSSES

The proposed method calculates three losses: two in Step 1, and one in Step 2. One of the losses used in Step 1 is the SFM_Loss. This factor represents the loss between the result of the FM module and the result of the resizing of the ground truth through the mean square error, as follows.

$$\text{SFM\_Loss} = \frac{1}{n}\sum_{p}(I_{p,FM} - I_{p,resized\_gt})^2 \qquad (3)$$

In eq. 3, p is a pixel, n is the total number of pixels, $I_{p,FM}$ is the output of the FM module, and $I_{p,resized\_gt}$ is the resized ground truth image. In Step 1, the filter selection

loss FS_Loss is also used. FS_Loss is calculated as the loss between the results of FS_GT and FS as follows.

$$\text{FS\_Loss} = \frac{1}{N}\sum_{k}\left|F_{k,FS} - F_{k,FS\_GT}\right| \tag{4}$$

k is a filter, N is the total number of filters, F represents the output of the FS module and the FS_GT module. Since the results of the FS_GT module and the FS module are either 0 or 1 for both filters, the L1 loss function, which is easier to calculate than the L2 loss function, is used. The reason for using two loss functions in Step 1 is that using only SFM_Loss cannot induce the network to select a desired filter, causing in a similar result to that of the segmentation network alone. The Seg_Loss used in Step 2 is calculated by the mean square error between the segmentation network result and the ground truth. This calculation is the same as that of the SFM_Loss of Step 1, but it has a different size.

III. EXPERIMENTAL RESULTS

A. TEST CONDITIONS

We used Pytorch in Ubuntu 16.04, GPU NVIDIA GeForce GTX 1080, CPU intel core i7-4770K and 16GB RAM. The size of the image is 256×256. Two hundred images were used for training, and 1787 images for testing. To achieve generalization, more data were allocated to the test set and fewer data were used for training. Other techniques were applied to supplement the size of the training dataset. With elastic deformation [14], the number of training images increased from 200 to 2000. In addition, vertical flip, horizontal flip, and 90°, 180°, and 270° rotation were randomly applied to training data [14]. A dataset was prepared from X-ray coronary angiography of 1180 patients who visited Asan Medical Center between September 2003 and July 2017. Two experts with more than five years of experience split the left anterior descending artery from the ostium to the distal site by using The CAAS QCA system (Pie Medical Imaging BV, the Netherlands) [15].

B. PERFORMANCE EVALUATION METHODS

Five criteria—accuracy, precision, recall, specificity, and F1 score—are employed to compare segmentation performance. First of all, we calculate true positives (TP), false positive (FP), true negatives (TN), and false negatives (FN) for each image. The value of a pixel corresponding to the blood vessel is 1 and the other value is 0 in the ground truth image and the result image. TP and TN thus mean the number of pixels, 1 or 0, respectively, both in the ground truth and in the result. FP means the number of pixels where the ground truth is 0 and the result is 1. FN means the number of pixels with a ground truth of 1 and a result of 0. (Table 1.) The five comparison criteria can be calculated using TP, FP, TN, and FN as follows:

$$\text{Accuracy(ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$\text{Precision(PRC)} = \frac{TP}{TP + FP} \tag{6}$$

TABLE 1. Definition of True positive (TP), True negative (TN), False positive (FP), False negative (FN).

		Ground truth	
		Vessel(1)	Background(0)
Test result	Vessel(1)	TP	FP
	Background(0)	FN	TN

$$\text{Recall(RC)} = \frac{TP}{TP + FN} \tag{7}$$

$$\text{Specificity(SPC)} = \frac{TN}{TN + FP} \tag{8}$$

$$\text{F1score(F1)} = 2 \cdot \frac{PRC \cdot RC}{PRC + RC} \tag{9}$$

The five comparison reference values all range from 0 to 1, with a higher value indicating better performance.

C. RESULTS

To demonstrate the effectiveness of SFM, we used the same pre-processing method and original U-Net as a segmentation network to compare the performance with the original U-Net. Firstly, to decide upon the pre-processing method, we applied three pre-processing methods: global histogram equalization [21]; local histogram equalization [21]; and Frangi filter to the original U-Net, and compared their performances (Table 2). Since the performance of the LAD main vessel segmentation was the best when the Frangi filter was used, the Frangi filter was applied to the proposed method and the original U-Net.

TABLE 2. The results of LAD main vessel segmentation using original U-Net with pre-processing methods.

Pre-processing methods	Sensitivity	Specificity
None	**0.403**	0.972
Global Histogram Equalization	**0.442**	0.973
Local Histogram Equalization	**0.346**	0.995
Frangi Filter	**0.495**	0.993

In addition, as mentioned above, the proposed method can use various existing networks as the network for SFM. We compared the experimental results of three network structures for SFM. The three networks are the contracting path of the U-Net [16], VGG16 [17] with batch normalization, and DenseNet121 [18] without transition layer three and dense block four. The classification layer is not used in either VGG16 or DenseNet121; the detailed structure is shown in Table 3. Prior to the comparison of performance, we briefly introduce the U-Net, VGG and DenseNet algorithms.

U-Net uses a networking and training method with a small amount of data. The CNN structure is divided into a contracting path and an expanding path. The two paths are symmetrical to each other and are connected by a bridge. The contracting path is used for capturing the context, and the expanding path is for accurate localization. Through the use

TABLE 3. Three kinds of architecture used in the network of SFM.

Output size	U-Net Contracting path	VGG16	DenseNet121
128x128	[3x3 conv] x2 2x2 MaxPool	[3x3 conv] x2 2x2 MaxPool	7x7 conv, stride 2
64x64	[3x3 conv] x2 2x2 MaxPool	[3x3 conv] x2 2x2 MaxPool	3x3 MaxPool, stride 2
32x32	[3x3 conv] x2 2x2 MaxPool	[3x3 conv] x3 2x2 MaxPool	[1x1 conv, 3x3 conv] x6 1x1 conv 2x2 AvgPool, stride 2
16x16	[3x3 conv] x2 2x2 MaxPool	[3x3 conv] x3 2x2 MaxPool	[1x1 conv, 3x3 conv] x12 1x1 conv 2x2 AvgPool, stride 2
16x16	[3x3 conv] x2	[3x3 conv] x3	[1x1 conv, 3x3 conv] x24 1x1 conv

of these paths, end-to-end training with very few images shows better performance in neural tissue segmentation than a sliding window convolutional network [19]. U-Net is well known as being the most widely used and effective architecture for medical image segmentation [5]. Thus, U-Net is used as a benchmark for performance comparison of the segmentation networks in this paper.

VGG took first and second place in the localization and classification tracks of ImageNet Challenge 2014. It was designed to improve the accuracy of the depth of a convolution network, and to evaluate the depth of the network using a very small (3×3) convolution filter. The network may consist of 11, 13, 16, or 19 layers, of which 16 and 19 showed the best performance. In this paper, we use VGG16, which has fewer parameters than the 19 layer network, but still had one of the best performances.

DenseNet is a method of connecting each layer to all other layers in a feed-forward manner. The algorithm is based on the observation that convolution networks are much deeper, more accurate, and more efficient when the number of connections between nearest input and near output layers is small. The U-Net contracting path and VGG16 layer are connected only to the previous layers, but DenseNet connects each layer to all previous layers. In this work, we used a DenseNet with 121 layers, which has a similar Top 5 error rate to VGG16.

The performance of the SFM is evaluated based on recall, which is the ratio of pixels included in the candidate area to the pixels of the main vessel in the ground truth image, since the purpose of SFM is to detect the area having the LAD main vessel with a bounding box. As shown in Table 4, the recall of SFM with U-Net, VGG, and DenseNet is 0.847, 0.843 and 0.834, respectively, and the average is 0.841. In other words, about 84% of the pixels of the main vessel are detected in the SFM step. In addition, the detection ratio of SFM with U-Net, VGG, and DenseNet is 96.36%, 95.97%, and 96.03%, respectively, based on a recall of greater then or equal to 0.5.

TABLE 4. The result of the SFM step.

	SFM(U-Net)	SFM(VGG)	SFM(DenseNet)
Recall	0.847	0.843	0.834
Detection ratio (%)	96.36	95.97	96.03

TABLE 5. Segmentation results of U-Net and the proposed methods with various network according to frangi parameter. (a) original U-Net. (b) SFM (U-Net). (c) SFM (VGG). (d) SFM (DenseNet). (e) SFM (Average).

(a)

Frangi parameter		PRC	RC	SPEC	F1	ACC
β	c					
0.5	15	0.719	0.441	0.995	0.519	0.979
1.0	0.1	0.698	0.510	0.994	**0.565**	**0.980**

(b)

Frangi parameter		PRC	RC	SPEC	F1	ACC
β	c					
0.5	15	0.800	0.607	0.996	**0.676**	**0.985**
1.0	0.1	0.621	0.616	0.989	0.603	0.978

(c)

Frangi parameter		PRC	RC	SPEC	F1	ACC
β	c					
0.5	15	0.767	0.609	0.995	**0.666**	**0.984**
1.0	0.1	0.583	0.665	0.986	0.608	0.976

(d)

Frangi parameter		PRC	RC	SPEC	F1	ACC
β	c					
0.5	15	0.724	0.587	0.993	**0.635**	**0.982**
1.0	0.1	0.594	0.603	0.989	0.585	0.977

(e)

Frangi parameter		PRC	RC	SPEC	F1	ACC
β	c					
0.5	15	0.764	0.601	0.995	**0.659**	**0.984**
1.0	0.1	0.599	0.628	0.988	0.599	0.977

Table 5 shows the performance of the original U-Net, SFM with three networks—U-Net, VGG, and DenseNet—and the SFM average using Frangi parameters. The Frangi parameters were determined experimentally, with appropriate c values based on β. Based on the F1 score, U-Net shows better performance when β is high and c is low. That is, the more clearly the blood vessels are detected in Fig. 5, the better the performance. Conversely, the proposed method gives better performance when β is low and c is high. In other words, the more the surrounding background is filtered in Fig. 5, the better the performance. Also, when comparing the results of SFM with three different networks, the performance of SFM with DenseNet121 is lower than those of U-Net's contracting path and VGG16. This is because all preceding layers are directly connected to the last layer in DenseNet121, unlike the other two networks. Similarly, experiments using SFM on the bridge of the original U-Net did not improve performance. This is because SFM is applied only to the last convolutional layer, thus the effect of the preceding layers on the last

layer is to reduce the effect of SFM. Comparing the average highest performance of SFM (Table 5 (E)) and original U-Net (Table 5 (A)), SFM shows higher performance—precision 0.066, recall 0.091, specificity 0.001, F1 score 0.094 and accuracy 0.004—than the original U-Net.

For a more detailed performance comparison, F1 score and accuracy were compared in several ways against the best performance of the original U-Net and SFM. Table 6 shows the mean and median values of the F1 score and accuracy of the original U-Net and SFM. On average, the median of the F1 score is 0.101 higher than the original U-Net, and the accuracy is 0.005 higher. Fig. 8 shows the density plot of F1 score and accuracy in the test data set. As shown in the figure, the SFM graph shows that the F1 score and the accuracy are both higher than that of the original U-Net. Also, Similar to the results in Table 6, SFM with DenseNet shows lower performance than the other two networks. We also applied a Kruskal-Wallis test [22] to compare the F1 score and accuracy between the original U-Net and SFM. Kruskal-Wallis is a nonparametric method used to analyze data that do not satisfy the normality and homogeneity assumptions of variance, and has a value between 0 and 1. If the value of the Kruskal-Wallis test is less than 0.05, it is judged that there is a difference between the values of the comparison objects. When comparing the Kruskal-Wallis value of the F1 score and accuracy between the original U-Net and the SFM average, both values are below 0.05, indicating that there is an important difference between the performance of the two algorithms. We also obtained Kruskal-Wallis values between two of the four methods, the original U-Net and SFM, with three networks. The Kruskal-Wallis value of the F1 score between SFM with U-Net and SFM with VGG is 0.300, which is more than 0.05, so we assume that there is no difference between the two algorithms. Since the Kruskal-Wallis value is 0.05 or less between the other algorithms, there appears to be a difference between their performances.

TABLE 6. Mean and median value of f1 score and accuracy of original U-Net and sfm. (a) F1 score. (b) Accuracy.

(a)

	Original U-Net	SFM (Avg)	SFM (U-Net)	SFM (VGG)	SFM (Dense)
Mean	**0.565**	**0.659**	0.676	0.666	0.635
Median	**0.608**	**0.709**	0.720	0.722	0.684

(b)

	Original U-Net	SFM (Avg)	SFM (U-Net)	SFM (VGG)	SFM (Dense)
Mean	**0.980**	**0.984**	0.985	0.984	0.982
Median	**0.981**	**0.986**	0.987	0.987	0.984

When we examine the analysis results of F1 score and accuracy, it can be seen that there is little difference in accuracy compared with the F1 score as a whole. We use accuracy to mean the number of TPs and TNs in the total pixels, as shown in eq. 5. Therefore, there is not a large difference in the accuracy of the image when the number

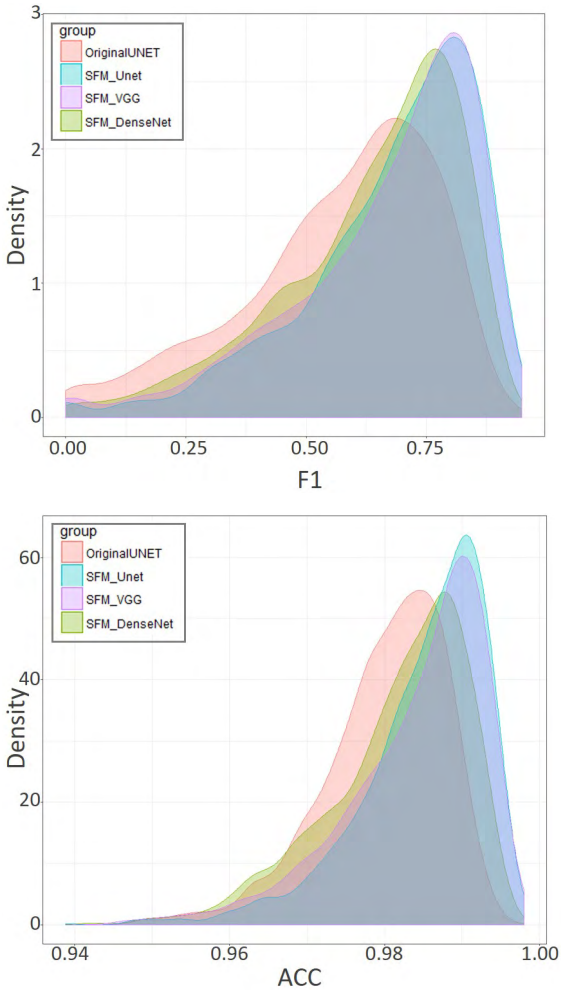

FIGURE 8. Density plot of F1 score (top) and accuracy (bottom) in test set.

of background pixels is too large compared to the number of pixels of the target object. Thus, we only calculated the correlation coefficient (r) of the F1 scores of the two methods and found it to be 0.79 (eq. 10).

$$\mathrm{r} = \frac{\sum (F_U - \mu_{F.U}) \cdot (F_{SFM} - \mu_{F.SFM})}{\sqrt{\sum (F_U - \mu_{F.U})^2 \cdot \sum (F_{SFM} - \mu_{F.SFM})^2}} \quad (10)$$

F_U is the F1 score of the original U-Net, and F_{SFM} is the F1 score of the average of SFM over test images $\mu_{F,U}$ is the mean value of the F1 scores of the original U-Net, and $\mu_{F,SFM}$ represents the mean value of the F1 scores of the average of SFM over the test images. Since the correlation coefficient is greater than zero, the two have a positive correlation. Also, the graph is located at the bottom of y = x, so SFM has better overall performance than the original U-Net (Fig. 9). At the bottom of the graph, images with a value of zero in the original U-Net axis have higher values on the SFM axis. In other words, images that do not have any pixels of the LAD main vessel by original U-Net can be detected better by SFM.

Fig. 10 shows the result of segmentation of the original U-Net and SFM with three networks: the contracting path of

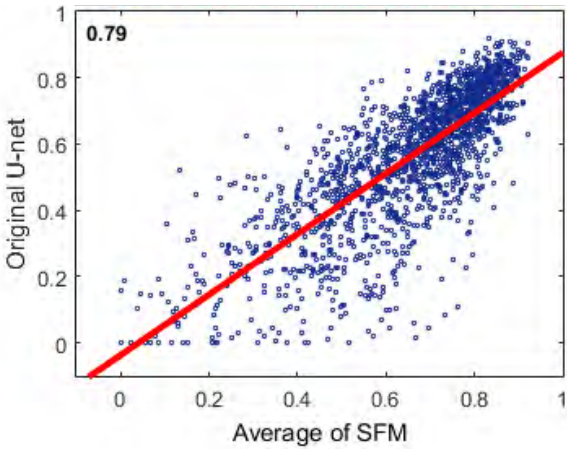

FIGURE 9. Correlation of F1 score between original U-Net and the average of SFM (Blue dot: F1 score of each test image, red line: graph of correlation based on blue dots).

U-Net, VGG16, and DenseNet121. The first column is the angiogram, which is the original image, and the last column is the segmentation (shown in red) of the main LAD vessel area of ground truth onto the original image. As shown in the figure, the proposed method, SFM, detects the position of the LAD main vessel in an angiogram, but segmentation results in some pixels being lost.

Fig. 11 shows a segmentation error. In the case of overlapping blood vessels, the proposed method judges both vessels as the main vessels of the LAD during segmentation, whereas the original U-Net does not segment all of the vessels. That is, if the judgment is ambiguous, the proposed method judges all of them as part of the LAD main vessel, but the original U-Net chooses to discard all of them. Also, if the image is too blurry or there is a catheter, the main vessel may not be properly detected or the catheter may be misunderstood as part of the main vessel. This phenomenon is similar to those seen when using the original U-Net, and as a result segmentation error increases. From these experimental results we conclude that the proposed method detects the main blood vessel area of LAD better than U-Net.

Therefore, rather than using the angiography image directly as an input, the region of the LAD main vessel is first detected from the angiography image and the result is used as input to the segmentation network in order to better detect the LAD main vessel. The reason for the poor performance when using an angiography image as input to U-Net is shown in Fig. 12. After the first epoch, even though there is some noise, all blood vessels are well segmented. As training progresses, however, the thinner part of the main vessel gradually disappears. From this it can be seen that the CNN is learned to discard thin blood vessels, even when the CNN is trained for segmentation using a ground truth consisting only of major blood vessels.

To solve this problem, the proposed method adopts a method of combining only filters with high detection rates of major blood vessels. Unlike the original U-Net, the proposed

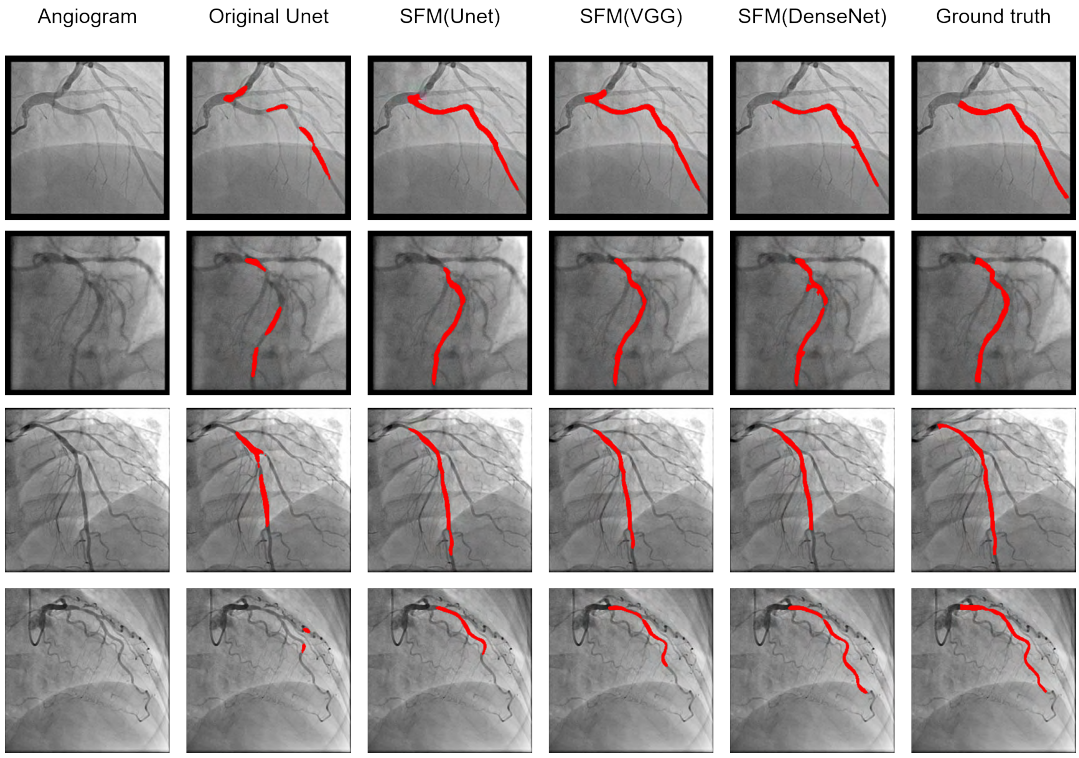

FIGURE 10. Segmentation results of SFM (with U-Net contracting path, VGG16, and DenseNet121) and original U-Net (each row represents each image).

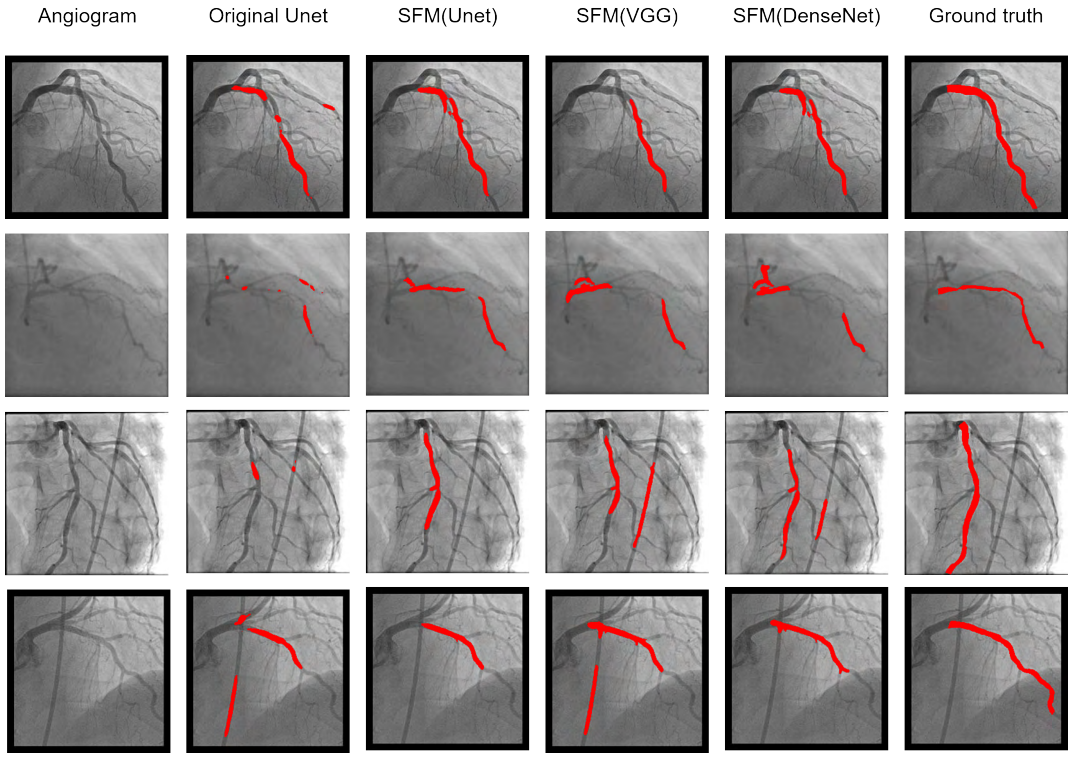

FIGURE 11. Segmentation error results of SFM (with U-Net contracting path, VGG16, and DenseNet121) and original U-Net (each row represents each image).

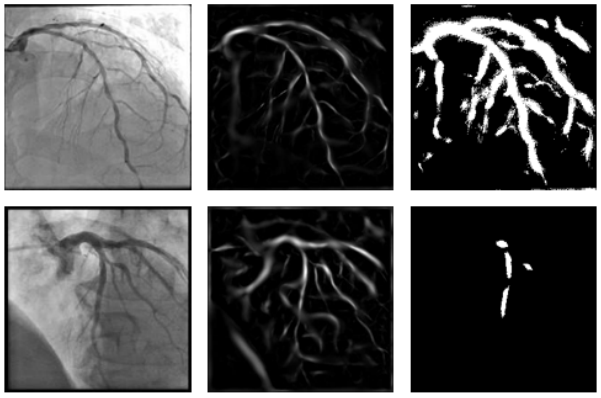

FIGURE 12. Training results of U-Net (top rows: after 1st epoch, bottom rows: after finishing training, left: angiography image, middle: pre-processed image, right: result image).

method uses the candidate area of the LAD main vessel for training, so it can reduce the side effects caused by removing the surrounding blood vessels.

IV. CONCLUSION AND FUTURE WORK

In this paper, we propose a selective feature mapping method for segmenting the main blood vessels of LAD in coronary angiography. Our algorithm, SFM, automatically segments a specific vessel, the LAD main vessel, from an image with multiple vessels. For this purpose, the network automatically selects feature maps that are concerned with main vessels. This algorithm can be applied to detect a specific object in an image in which several different objects have been classified into the same category. In addition, unlike previous studies that require the creation of the ground truth of all blood vessels for the detection of a particular blood vessel, it requires less effort to generate the dataset because it uses only the ground truth of the one vessel of interest.

According to the experimental results, the proposed method showed higher performance than the original U-Net, which is popular in medical image segmentation, in all five comparison criteria. More specifically, the proposed method has a precision of about 0.066, recall of 0.091, specificity of 0.001, F1 score of 0.094 and accuracy of 0.004 higher than the original U-Net. These results come from comparing the average of the three networks used in the proposed method and the performance of the original U-Net. The reason for the low performance of the original U-Net is that the original U-Net is trained to discard the peripheral vessels in order to segment the main vessel from an angiogram, resulting in a side effect where part of the main vessel is discarded. In order to compensate for this disadvantage of the original U-Net, the proposed method first detects the area of the LAD main vessel using SFM and uses it for training. For this purpose, SFM was tested with three networks, and there is a slight difference in performance depending on the network used. Therefore, it is expected that better results will be obtained if the results of the three networks are selected competitively. Also, the performance of the proposed method depends on

the detection accuracy of the candidate region in the first step. Therefore, it is necessary to apply more diverse networks to detect candidate areas.

The proposed method selects a specified number of feature maps. If the network structure is redesigned so that the CNN can select the appropriate parameters for filters, it will be able to select only the optimal filters according to the characteristics of each image, and so should perform better.

In addition, if a catheter is present, it is recognized as the main blood vessel, and as a result, detection performance deteriorates. Therefore, if a catheter is detected at the beginning of the image analysis, and information about the catheter is presented to the CNN, segmentation performance can be improved by allowing the network to automatically remove pixels representing the catheter. One way to achieve this removal is to filter the feature maps that detect the catheter, as well as those that detect the main vessel. However, in this case, the need for ground truth regarding the catheter may require the creation of additional data sets for filtering.

Finally, the proposed method shows higher performance than the original U-Net, but it shows inadequate performance compared to the ground truth produced by CAAS QCA. To illustrate the process of segmenting the main vessel of the LAD using CAAS QCA, an expert inputs a DICOM image into CAAS and selects the image that best shows the LAD main vessel among the input DICOM images. The specialist then clicks on five or six points on the image, they are connected to form a single line, and the CAAS system automatically segments the blood vessel based on this line. The expert then corrects any errors in the segmented blood vessel. CAAS QCA has obvious advantages in that the vessel segmentation is performed with fairly high accuracy in a short time. However, it does not generate objective, uniform results, as it requires expert input. There is also a limitation in that the CAAS system can not locate the main vessel of the LAD. In contrast, the proposed method can generate uniform results and locate the LAD main vessel autonomously. Instead of finding the approximate location of the LAD main vessel through an expert's click, a skeleton of the final result of the proposed method or a candidate area can be used as input to the CAAS. In this case, it is possible to overcome the limitations of the CAAS system, which performs semi-automatic segmentation.

REFERENCES

[1] H. R. Fazlali *et al.*, "Vessel region detection in coronary X-ray angiograms," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 1493–1497.

[2] I. Cruz-Aceves, F. Oloumi, R. M. Rangayyan, J. G. Aviña-Cervantes, and A. Hernandez-Aguirre, "Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization," *Biomed. Signal Process. Control*, vol. 25, pp. 76–85, Mar. 2016.

[3] E. Nasr-Esfahani *et al.*, "Vessel extraction in X-ray angiograms using deep learning," in *Proc. IEEE 38th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 643–646.

[4] H. R. Fazlali *et al.*, "Vessel segmentation and catheter detection in X-ray angiograms using superpixels," *Med. Biol. Eng. Comput.*, vol. 56, no. 9, pp. 1515–1530, 2018.

[5] A. Vlontzos and K. Mikolajczyk. (2018). "Deep segmentation and registration in X-ray angiography video." [Online]. Available: https://arxiv.org/abs/1805.06406

[6] E. Nasr-Esfahani *et al.*, "Segmentation of vessels in angiograms using convolutional neural networks," *Biomed. Signal Process. Control*, vol. 40, pp. 240–251, Feb. 2018.

[7] C. Kondermann, D. Kondermann, and M. Yan, "Blood vessel classification into arteries and veins in retinal images," in *Proc. SPIE*, vol. 6512, p. 651247, Mar. 2007.

[8] R. A. Welikala *et al.*, "Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort," *Comput. Biol. Med.*, vol. 90, pp. 23–32, Nov. 2017.

[9] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 1998.

[10] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson. (2015). "Understanding neural networks through deep visualization." [Online]. Available: https://arxiv.org/abs/1506.06579

[11] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2921–2929.

[12] M. Lin, Q. Chen, and S. Yan. (2013). "Network in network." [Online]. Available: https://arxiv.org/abs/1312.4400

[13] S. Ioffe and C. Szegedy. (2015). "Batch normalization: Accelerating deep network training by reducing internal covariate shift." [Online]. Available: https://arxiv.org/abs/1502.03167

[14] L. Perez and J. Wang. (2017). "The effectiveness of data augmentation in image classification using deep learning." [Online]. Available: https://arxiv.org/abs/1712.04621

[15] J. Haase *et al.*, "Experimental validation of geometric and densitometric coronary measurements on the new generation Cardiovascular Angiography Analysis System (CAAS II)," *Catheterization Cardiovascular Diagnosis*, vol. 30, no. 2, pp. 104–114, 1993.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015.

[17] K. Simonyan and A. Zisserman. (2014). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556

[18] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. CVPR*, vol. 1, no. 2, 2017, pp. 4700–4708.

[19] D. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 2843–2851.

[20] H. Yao, W. Zhang, R. Malhan, J. Gryak, and K. Najarian, "Filter-pruned 3D convolutional neural network for drowsiness detection," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 1258–1262.

[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, MA, USA: Addison-Wesley, 1992.

[22] P. E. McKight and J. Najab, "Kruskal-wallis test," in *The Corsini Encyclopedia of Psychology*. Hoboken, NJ, USA: Wiley, 2010, p. 1.

KYUNGMIN JO received the B.S. degree in electronic and electrical engineering from Hongik University, Seoul, South Korea, in 2006, and the M.S. degree in electrical and computer engineering from Sungkyunkwan University, Suwon, South Korea, in 2011.

From 2005 to 2009, she was a Researcher with Samsung Electronics Co., Ltd. Since 2015, she has been a Researcher with the Medical Engineering Research and Development Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul. Her research interests include computer vision, image processing, and artificial intelligence in medical applications.

JIHOON KWEON (M'17) received the B.S. degree in mechanical and aerospace engineering and the Ph.D. degree (Master's and Doctoral Integration Course) in mechanical and aerospace engineering from Seoul National University, Seoul, South Korea, in 2004 and 2011, respectively.

From 2011 to 2013, he was a Senior Researcher with the Institute of Advanced Machines and Design, Seoul National University. Since 2014, he has been a Research Assistant Professor with the University of Ulsan College of Medicine, Seoul. Since 2017, he has been a Research Assistant Professor with the Biomedical Engineering Research Center, Asan Medical Center, Seoul. His research interests include the development of diagnostic method for cardiovascular diseases, hemodynamic simulation, and medical image processing with machine learning.

Dr. Kweon is a member of the Korean Society of Mechanical Engineers.

YOUNG-HAK KIM received the B.S. degree in medicine from Kyungpook National University, Daegu, South Korea, in 1993, the M.S. degree in medicine from the University of Ulsan College of Medicine, Seoul, South Korea, in 1998, and the Ph.D. degree in medicine from Kyungpook National University, in 2005.

From 2003 to 2008, he was an Assistant Professor of medicine with the Asan Medical Center, University of Ulsan College of Medicine, where he was also an Associate Professor of medicine, from 2008 to 2014, and has been a Professor of medicine, since 2014. Since 2017, he has been the Director of the Health Innovation Big Data Center, Asan Medical Center. His research interests include the development of diagnostic method and medical device for cardiovascular diseases, and medical big data.

Dr. Kim is an Associate Editor of the *Korean Journal of Internal Medicine* and the *International Journal of Cardiovascular Imaging*.

JAESOON CHOI (M'96) received the B.S. degree in control and instrumentation engineering and the M.S. and Ph.D. degrees in biomedical engineering from Seoul National University, Seoul, South Korea, in 1995, 1997, and 2003, respectively.

He had predoctoral training at the Department of Biomedical Engineering, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH, USA, from 1999 to 2000. From 2004 to 2006, he was a Staff Researcher with the Research Institute, National Cancer Center, Seoul. From 2007 to 2012, he was a Research Professor with the Korea Artificial Organ Center, College of Medicine, Korea University, Seoul. He is currently an Associate Professor with the Department of Biomedical Engineering, Asan Medical Center, University of Ulsan College of Medicine, Seoul. His research interests include computer-aided surgery and intervention, and mechatronics system application in biomedicine.