



Research Paper

Segmentation of vessels in angiograms using convolutional neural networks



E. Nasr-Esfahani^a, N. Karimi^{a,*}, M.H. Jafari^a, S.M.R. Soroushmehr^{b,c}, S. Samavi^{a,b,c}, B.K. Nallamothu^c, K. Najarian^{b,c}

^a Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran

^b Michigan Center for Integrative Research in Critical Care, Ann Arbor, MI 48109, USA

^c Department of Emergency Medicine, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:

Received 15 January 2017

Received in revised form 28 August 2017

Accepted 16 September 2017

Keywords:

Angiograms

Vessel segmentation

Deep learning

Convolutional neural networks

ABSTRACT

Coronary artery disease (CAD) is the most common type of heart disease and it is the leading cause of death in most parts of the world. About fifty percent of all middle-aged men and thirty percent of all middle-aged women in North America develop some type of CAD. The main tool for diagnosis of CAD is the X-ray angiography. Usually these images lack high quality and they contain noise. Accurate segmentation of vessels in these images could help physicians in accurate CAD diagnosis. Many image processing techniques have been used by researchers for vessel segmentation but achieving high accuracy is still a challenge in this regard. In this paper a method for detecting vessel regions in angiography images is proposed which is based on deep learning approach using convolutional neural networks (CNN). The intended angiogram is first processed to enhance the image quality. Then a patch around each pixel is fed into a trained CNN to determine whether the pixel is of vessel or background regions. Different elements of the proposed method, including the image enhancement method, the architecture of the CNN, and the training procedure of the CNN, all lead to a highly accurate mechanism. Experiments performed on angiograms of a dataset show that the proposed algorithm has a Dice score of 81.51 and an accuracy of 97.93. Results of the proposed algorithm show its superiority in extraction of vessel regions in comparison to state of the art methods.

© 2017 Published by Elsevier Ltd.

1. Introduction

Coronary artery disease (CAD) is currently the number one cause of death in many parts of the world, causing one in every four deaths in the US and affecting more than 13 million people there. Healthy arteries are smooth and elastic. CAD occurs when arteries get narrowed or blocked because of fat, cholesterol or other substances, called plaque, circulating in the blood. Plaques that accumulate on the inner walls of arteries create stenosis. Stenosis in arteries causes restriction of the flow of blood to the heart and leads to critical situations such as heart attacks [1].

Among various methods that exist for diagnosing CAD disease, X-ray angiography is known to be the main standard procedure. In this method, the arteries are filled with a contrast agent in order to show the arteries in the X-ray images. For this purpose, a long and thin hallow tube, called catheter, is inserted into an artery in

the patient's arm or groin. The tube is then guided through the major coronary arteries that are intended to be examined. When the catheter arrives at the target arteries, the contrast agent is injected and a number of X-ray images are recorded. The taken sets of images, called angiograms, would visualize the coronary vasculature.

However, the captured angiograms usually have a low quality due to the existing limitations on X-ray radiation. The emission of X-ray in angiography is restricted due to its side effects on human body. As a result, usually there is a low contrast between the vessel region and its background. Various computer based methods are introduced in order to aid the usability of these images. Meanwhile, a critical step towards the detection of stenosis is the accurate segmentation of vessel regions from its background. The captured images may suffer from uneven brightness, presence of tissues other than arteries, noise and blurriness caused by heart motions and movements of the camera [2]. These problems make the segmentation of vessels in X-ray angiograms a challenging task.

Existing vessel segmentations can be categorized in three main groups of model-based, tracking-based, and pattern recognition-

* Corresponding author.

E-mail address: nader.karimi@cc.iut.ac.ir (N. Karimi).

based methods [3]. Active contours are a group of segmentation methods that belong to the model-based methods. Active contour methods include two main approaches [4] of edge-based and region-based. In edge-based models local edge information is used as a criterion to stop the curve from evolving. The region-based models perform by using statistical information from the inside and outside of the curve and are less affected by the presence of noise. The work of [4] is an example of active contour models for vessel extraction. A local feature fitting function is proposed to detect the vessel in different sizes. Also, it is supposed to be resistant against the presence of noise, inhomogeneity and complexity of the background in the image. Nonetheless, active contours approaches usually fail in the handling of non-uniform illumination. Also, active contours are highly dependent on the initial selection of contour and they usually have high computational complexity.

The tracking-based segmentation methods use local information to detect the vessels. Tracking-based methods start from an automatically or manually selected initial point and follow the path that better matches the vessel profile to track the vessel edges or the vessel centerlines. By completely scanning the image, a connected region that locates vessel regions would be obtained. In Ref. [5] an iterative probabilistic tracking method is proposed for the segmentation of retinal vessels. Authors of Ref. [5] manually select a starting point to track the vessel using continuity in vessel gray levels. The tracking-based models could deal with vessel crossings and bifurcations and they provide geometrical features of the extracted vessel. The method proposed by Zhou Sh. et al. [6] is another example of tracking-based method that uses fuzzy inferring and tracking. Authors of Ref. [6] firstly enhance the images by applying Gabor and Hessian filter. The artery vessels are chosen from a group of candidates derived by a probabilistic tracking operator (PTO). A sun-operator, as vessel structure pattern detector (SPD), assigns the extracted parts of the vessel to bifurcation or crossing groups [6]. The tracking-based models might be misled when the resolution is low or complex image background exists.

Filtering methods such as Hessian-based filters, Flux-based filters [7,8] and non-linear diffusion can be categorized as pattern recognition-based approaches. Frangi et al. [9] proposes analysis of eigenvalues in Hessian matrix for detection of the vessel region. Work of Refs. [10] and [2] are other filtering methods that take advantage of decimation-free directional filter bank (DDFB), Hessian filter, and guided filtering, for enhancement of X-ray angiography images. In Ref. [11], Hessian filter is used for detection of coronary arteries region-of-interest (ROI). It then performs the segmentation using both the extracted ROI and measurements of a flux flow filter. Also, in the work of Ref. [12], that is a pattern recognition-based method, an algorithm is proposed for automatic vessel extraction using graph-cuts [13]. Authors of Ref. [12] use graph-cut that is based on geodesic paths and multistage edge maps. However, geodesic methods are usually sensitive to noise and complex vessel structures and background regions.

Recently deep learning methods have shown promising results in various image processing and medical imaging applications [14–19]. Convolutional neural networks (CNN) are among powerful deep learning methods that are suitable for image processing applications. While these networks are introduced more than two decades ago [20], their usage has increased rapidly in recent years. The recent growth of usage is due to the development of techniques that have facilitated the training of these networks [21]. CNNs consist of several convolutional and pooling layers which are usually followed by one or more fully connected layers. CNNs can be used as an effective tool for automatic feature extraction.

Deep convolutional neural network is configured for the detection of retinal vessel in Ref. [14]. In Refs. [15,16] the CNN is used for the classification of lung disease. Works of Refs. [17] and [18] are other researches that have considered CNNs for medical imaging

applications for the detection of brain tumors and the diagnosis of skin cancer. Moreover, Ref. [19] proposes a deep learning approach for segmentation of lesions from skin images. In Ref. [22] Nasr et al. present a deep learning method for segmentation of vessels in angiograms. In Ref. [22], a single path CNN is used based on relatively large sized windows, hence it can be considered that global patches are being used. The use of single source of information in Ref. [22] results in a relatively poor performance in accurate detection of the boundaries of vessels in some images. Accurate segmentation of boundaries is essential in diagnosis of stenosis.

In the proposed method in this paper the networks accuracy is improved in comparison with the preliminary version of the work presented in Ref. [22] using the following three novelties:

- Combination of the information that exist in local and global patches.
- Usage of the neighboring pixels' labels that are obtained from an initial segmentation map.
- Taking advantage of edges locations that exist in patches produced by a canny edge detector.

Here, the authors propose a deep learning method for extraction of vessel in X-ray angiograms. The proposed method performs accurate extraction of vessel boundary regions in challenging situations of complex backgrounds and the presence of noise. In this proposed method, initially the input image is preprocessed to enhance the contrast and to reduce the image noise. Two convolutional neural networks are designed. The first CNN uses local and global image patches to establish an initial segmentation probability map. The input of the second CNN includes the output of the first CNN as well as edge information from canny edge detector. The second CNN corrects the outcome of the first, especially in the challenging border regions. A training mechanism is devised to train the networks to correctly distinguish vessel border regions from those of background and inside vessel regions. Experimental results demonstrate that the proposed method outperforms existing angiogram segmentation algorithms in terms of vessel extraction accuracy.

The organization of the rest of this paper is as follows. In Section 2 the proposed method for detection of vessel regions is explained. Experiments done for evaluation of the proposed method are given in Section 3 and Section 4 concludes the paper.

2. Proposed method

The block diagram of the proposed method is shown in Fig. 1. The proposed method consists of four steps: a) contrast enhancing, b) first CNN stage, c) second CNN stage, and d) binary vessel map formation. These four stages are shown in Fig. 1. The first step consists of an image enhancing procedure; the result is fed to the first CNN stage. The first CNN stage uses both local and global features to form a vessel probability map. We also feed the enhanced image into an edge detector stage to form an edge map. The second CNN stage, as shown in Fig. 1(c), receives both the edge map and the vessel probability map to find an enhanced vessel probability map. In the fourth stage, shown in Fig. 1(d), a threshold is applied to the enhanced probability map to form a binary map. The largest connected component in the final binary map shows the segmented vessels. In sub-sections 2.1–2.4 different steps of the proposed method are discussed.

2.1. Contrast enhancing

The contrast enhancing procedure is directed towards improving the performance of the first CNN stage. Thus, the proposed

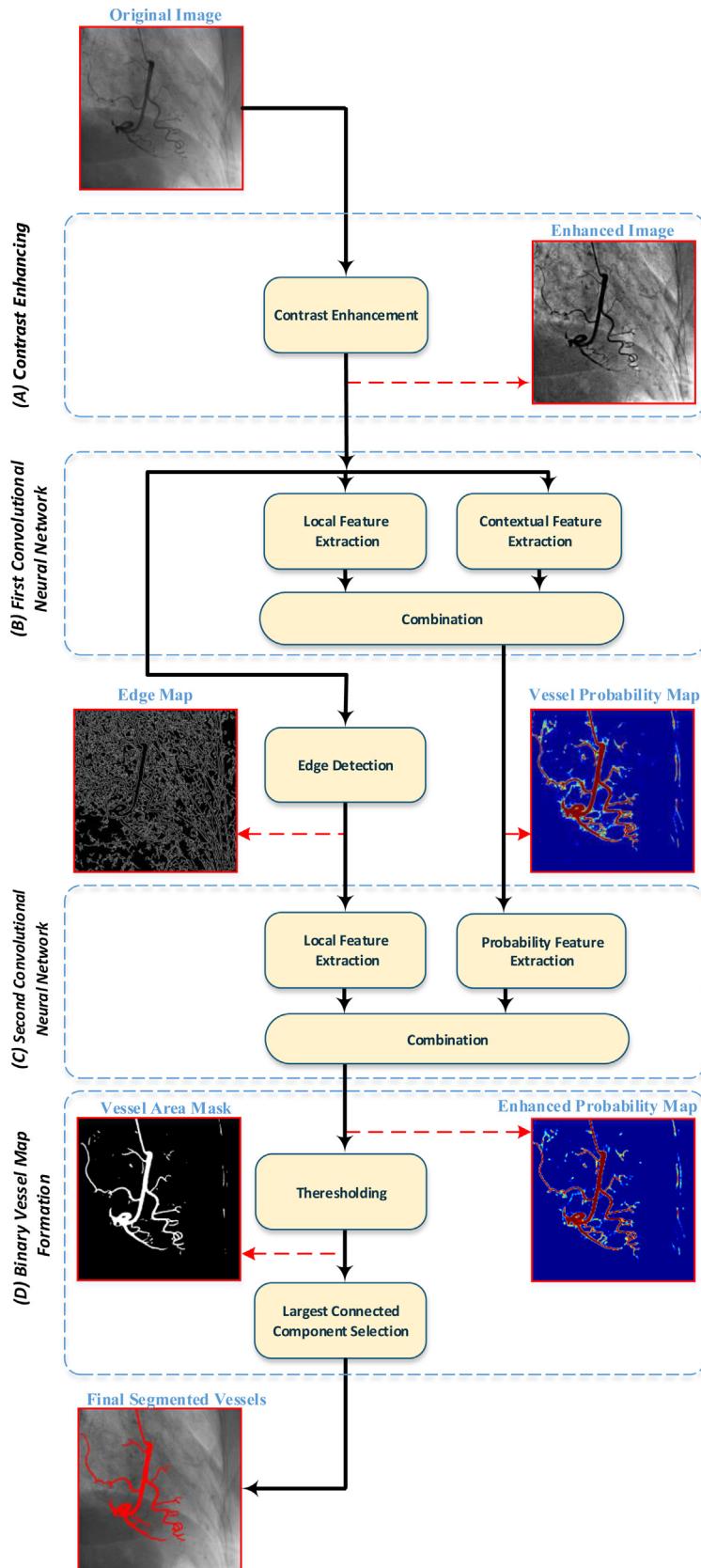


Fig. 1. Block diagram of the proposed method.

method involves applying a multi-scale top-hat transform [23] in order to enhance the contrast of input angiogram images. Top-hat transform is a mathematical morphological operation used as a

preprocessing or post processing stage in diverse image processing and pattern recognition applications [24,25]. Top-hat is based on morphological dilation and erosion operations. Two maps as

white top-hat and black top-hat are derived from the input image. In white top-hat the areas in the image that have high gray scale values, i.e. bright areas, are tracked and intensified. In contrast, in black top-hat the dark regions in the image are highlighted. Hence by applying Eq. (1), after the image is added by white top-hat and the black top-hat is subtracted from the image, the contrast of the image is increased.

$$f_{en} = f + f_w - f_b \quad (1)$$

In (1), f is the input and f_{en} is the output enhanced image. Also f_w and f_b denote the white and black top-hat maps respectively.

In multi scale top-hat transform, different scales are selected for structuring element B_i . Using Eqs. (2) and (3) to corresponding black and white top-hat maps are calculated for each scale.

$$f_{w_i} = f - ((f \ominus B_i) \oplus B_i) \quad (2)$$

$$f_{b_i} = ((f \oplus B_i) \ominus B_i) - f \quad (3)$$

In Eqs. (2) and (3), B_i , f , f_{w_i} and f_{b_i} denote the structuring element of the i^{th} scale, input image, the white and black top-hat transforms respectively. Also, \oplus , \ominus are morphological dilation and erosion operators. The dilation and erosion of $f(x, y)$ by $B_i(u, v)$ are defined by Eqs. (4) and (5).

$$f \oplus B_i = \max_{u,v} (f(x-u, y-v) + B_i(u, v)) \quad (4)$$

$$f \ominus B_i = \min_{u,v} (f(x+u, y+v) - B_i(u, v)) \quad (5)$$

Structuring element B_i is initialized with a small size (3×3) and then its size is increased until it reaches its maximum scale (19×19). Then for every position in the produced set of white top-hat maps the maximum value is considered to produce the final single white top-hat map. The same procedure is performed for the set of black top-hat maps. The two finalized maps are used in (1) to produce an enhanced image. Sample angiogram images preprocessed by multi scale top-hat transform are shown in Fig. 2.

The enhanced image is applied to a convolutional neural network with the architecture explained in Section 2.2. As can be seen in Fig. 1, the proposed algorithm consists of two CNN stages.

2.2. Architecture of the first CNN stage

In order to segment vessel regions from the background, the image is fed to a CNN by patches. The patches are defined using a window around a pixel. This window is slid on the entire image with stride of one. Hence a patch would be considered around each pixel in the image where the pixel is located in the center of the patch. The patches are applied to the CNN and the output shows the probability of that pixel in the center of the patch belonging to vessel region. As a result, by applying all of the patches, a probability map with the same size as the input image is obtained. This map shows the probability of each pixel belonging to a vessel region.

The architecture of the first CNN, called CNN1, is shown in Fig. 3. The network that is used in the proposed method comprises both, a Pooling and a Convolutional Layer. It should be noted that fully connected layers can be considered as a convolutional layer with kernel size equal to the size of input feature map.

Suppose that F_j^l is the j^{th} feature map in the l^{th} layer of the network. Also N^l is the number of the feature maps in the l^{th} layer. The output of the l^{th} convolutional layer is calculated by (6). In Eq. (6), w_{ij}^l denotes the i^{th} kernel used to produce the j^{th} feature map and b_j^l is the j^{th} bias in the l^{th} layer.

$$F_j^l = \sum_{i=1}^{N^l} \left(F_j^{l-1} * w_{ij}^l \right) + b_j^l \quad (6)$$

Moreover, the pooling layers are set to be max-pooling that return the maximum value in a window of size $S^l \times S^l$. The output of the l^{th} max-pooling layer is calculated by (7).

$$F_j^l(x, y) = \max \left\{ F_j^{l-1}(x+m, y+n); (m, n) \in \{0, \dots, S^l\} \right\} \quad (7)$$

There are two parallel paths in the CNN1 architecture of Fig. 3. One of the paths is called Contextual Feature Extraction path. This path consists of a total of 4 layers, which include 2 convolving layers. Each convolving layer is followed by a max pooling layer. The second path is the Local Feature Extraction path. Its three layers include a convolving layer which is followed by a max pooling layer and another convolving layer. One path is to process local patches and the other one performs on global patches. Around each pixel two different size patches are formed. The smaller patch is called local and the larger patch is the global patch. The global patch should give a global view and it basically contains contextual information. On the other hand, the local patch shows the detailed information of the local texture of the image. It should be noticed that each pair of local and global patches have the same central pixel. The outputs of CNN1 indicate the probability of each pixel belonging to a vessel region.

The global patch is a window with size 75×75 and the local patch is a 13×13 window. Global extracted patches are then resized to 25×25 . As can be seen in Fig. 3 the first layer in both global and local paths is a convolutional operator where patches are convolved with filters. There are 50 different filters with 6×6 kernel size for the global path. This would result in 50 patches with the size of 20×20 . Also, for the local path there are 50 filters with 4×4 kernel size that the input patch is convolved with them. This results in 50 patches of the size 10×10 . The first convolutional layers are followed by max-pooling layers with kernel size 2×2 . The third layer in both paths is again a convolutional layer with 20 filters with 5×5 kernel size in the global path and 20 filters with 3×3 kernel size in the local path. The second convolutional layer in the global path is followed by a max-pooling layer with 2×2 kernel size.

By passing the local and global patches through their corresponding paths, the local and global information around each pixel is processed. The extracted information is combined and fed to a fully connected network with 100 neurons in the input and 2 neurons in the output layer. The 2 output neurons show the probability that the pixel in the center of the two input patches belongs to a vessel region based on a softmax function. The softmax function is defined as (8).

$$p_k = \exp(a_k) / \sum_{k'=1}^C \exp(a_{k'}) \quad (8)$$

In (8), a_k and p_k denote output of neuron k and the probability of the input's membership to the k^{th} class, respectively. Also C is the number of classes. In the vessel segmentation case there are two classes of foreground and background. It was shown in Fig. 3 that the architecture of CNN1 has two paths that process local and global information separately. The extracted information is then combined in the last layer. The use of both local and global information is essential and can improve the performance of segmentation. In a Local window with a small size, the local edges of the vessels can be extracted. Meanwhile some artifacts may be wrongly detected as vessel regions. It is shown in Fig. 4(c) that if CNN1 were to use only local patches it would have made mistakes in classification of some local artifacts. Contrary to local patches, the global patch system would show the global context and shape of the vessels and also it is not affected by local artifacts. But relying only on global patches would result in inaccurate extraction of the vessels. An example of an output probability map of a CNN1, based only on global patches, is presented in Fig. 4(d). As can be seen, the

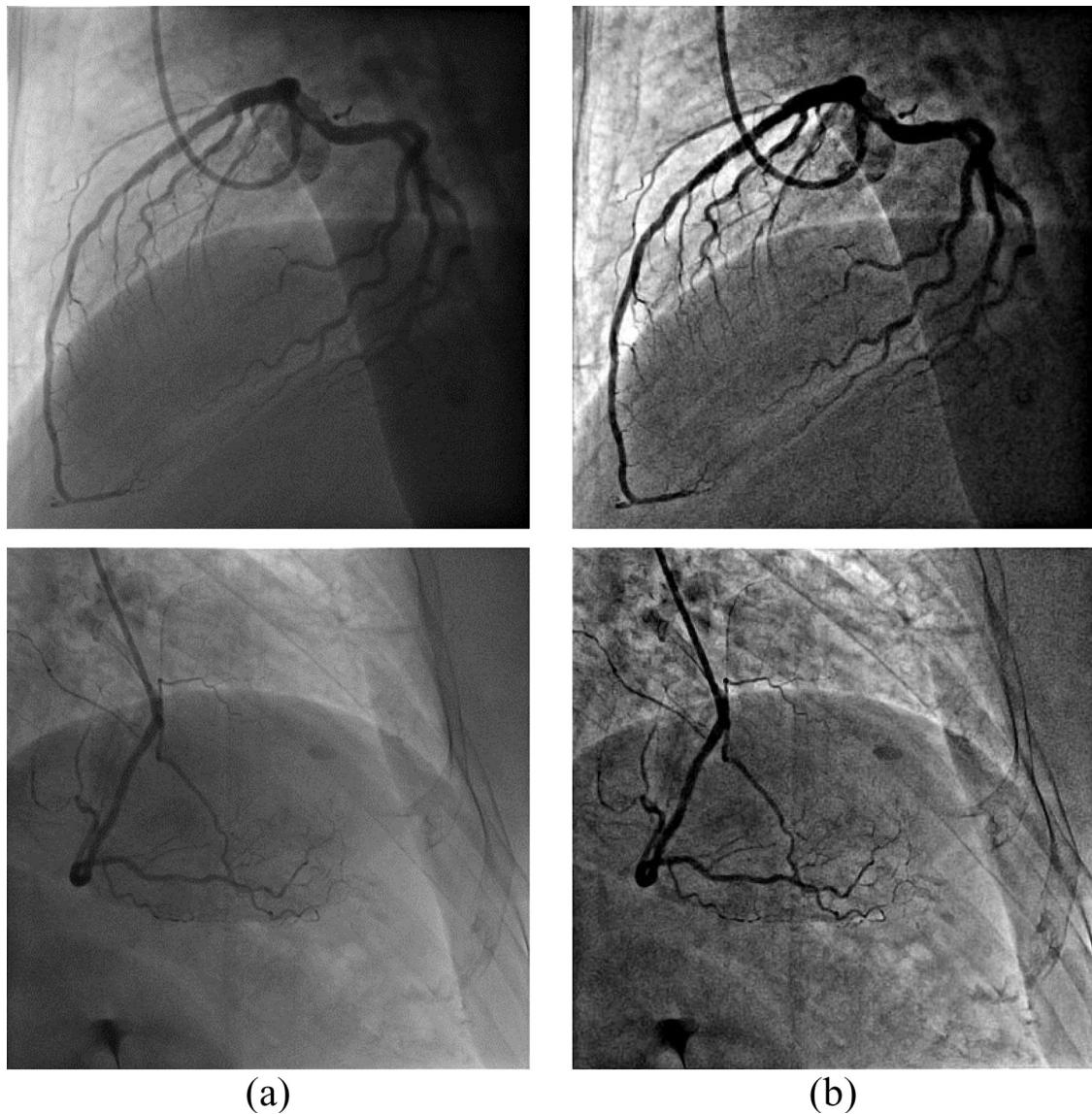


Fig. 2. Contrast enhancement, (a) input image, and (b) enhanced image by multi scale top-hat transform.

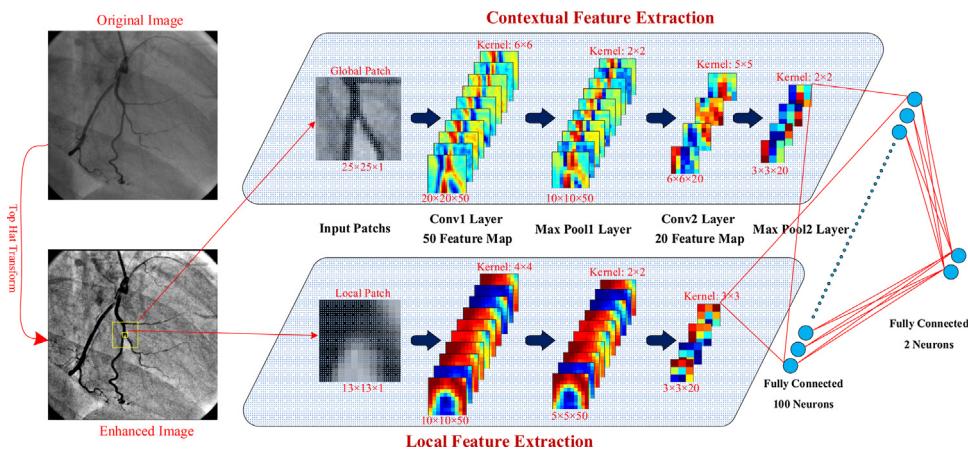


Fig. 3. Architecture of the first CNN stage (CNN1).

output of global network has detected the global structure of the vessel more accurately, while showing a poor performance in the edge regions. As a result, as shown in Fig. 4(e), by combination of

local and global information the accuracy of the system for border areas is enhanced while it is not distracted by local shapes and artifacts.

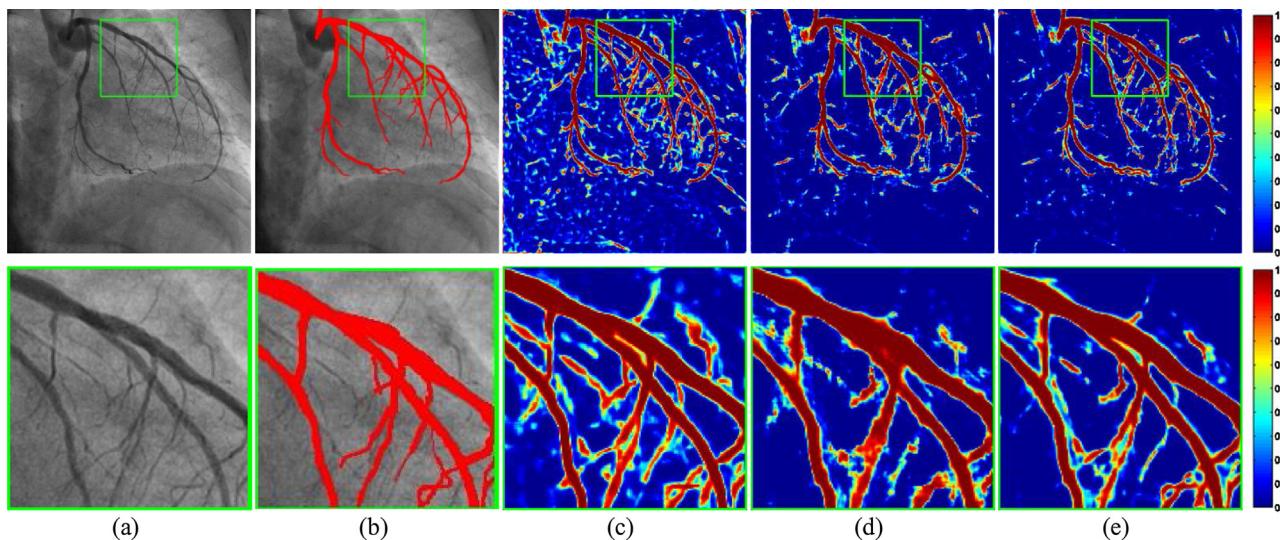


Fig. 4. Intermediary results, (a) input image, (b) ground truth segmentation, probability maps (c) based on local patches, (d) based on global patches, (e) based on both local and global patches.

2.3. Architecture of the second CNN stage (CNN2)

The output of the first stage could detect the vessel regions with a relatively high precision but it had some difficulty in the vessel border regions. Meanwhile the segmentation map of the first stage would be further refined by being applied to the second CNN stage, CNN2. The inputs to the second stage are the probability map obtained from CNN1, along with a map that shows the location of edges in the image.

Precise detection of vessel parts is a challenging task especially in the vessel border regions. Incorrect classification of pixels in the border of the vessels would result in a significant decreasing of the segmentation performance. In order to enhance the accuracy of the final decision in these challenging parts, in the second stage of this proposed method the location of edges are also fed to CNN2. Among various existing methods for edge detection [26], in this paper we apply canny edge detector as one of the most commonly used methods. In the canny method, firstly the image is filtered by a Gaussian filter to reduce noise effects. The gradient of the image is calculated and the probable locations of edges are extracted by non-maximum suppression technique. Weak edges are then eliminated by applying a double threshold and tracking the location and neighborhood of the extracted edge points.

In the proposed method the threshold for the canny edge detection is set to a low value; hence most of the edges in the image would appear in the map. The canny map is used to help the segmentation procedure by focusing on the edge regions. The second CNN is directed to notice these critical edge locations and hence refine the initial probability map of CNN1. In the second stage, the edge map and the initial probability map of CNN2 are processed through two parallel paths of CNN2. A window of size 13×13 is used to form a patch around each position in the initial probability map of CNN1. The same window size is used around each position in the canny edge map to form an edge patch. These two patches are fed to the two paths of CNN2. Then the output would show the refined probability for the corresponding pixel in the input image belonging to a vessel region.

The architecture of the second network and the configured layers are shown in Fig. 5. As can be seen, the two paths in the second stage have similar layers. These layers consist of convolutional layer1, max-pooling layer1, and convolutional layer2. The kernel size of the 50 filters of the first convolutional layer is 4×4 . There are 20 filters in the second convolutional layer where each filter

has a kernel of the size 3×3 . Hence, the first convolutional layer produces 50 feature maps and the second convolutional layer produces 20 feature maps. The size of kernel in the max-pooling layer is 2×2 . At the end, the two parallel paths are fed into a two layer fully connected layer with 100 neurons in first layer and 2 neurons in the output layer. This would provide the elements of a refined probability map.

The purpose of the second convolutional neural network, CNN2, is to enhance the accuracy of the initial probability map. The enhancement is done due to the available label for neighboring pixels from the initial probability map, along with the information about the location of the edges extracted by the canny operator. These improvements are shown in Fig. 6, where sample probability maps produced in the first and the second CNN stages are compared. As can be seen, the output of CNN2 has a better performance in edge regions. Moreover in the refined probability map of CNN2 the continuity of some vessels is preserved and some wrongly classified artifacts that had affected the segmentation map of CNN1 are eliminated.

It should be mentioned that the edge information could have been applied to the first CNN stage but inferior results are obtained as compared to the proposed two stage structure. The accuracy of a single stage CNN even when edge information is utilized is less than the obtained results of the proposed method. This is due to the fact that the second CNN also has the probability labels of the neighboring pixels when it uses the edge information to decide on the final probability label of each pixel. But if the proposed method were to use a single CNN stage each pixel is processed independently without the knowledge of the labels of its neighboring pixels. This improvement was shown in the conducted experiments. It was shown that addition of edge patches in a single stage CNN structure would increase the performance of the CNN by one percent in terms of Dice score. While utilization of two stage CNN system and addition of edge patches to the second CNN would increase the Dice score by 4%.

2.4. Binary vessel map formation

As the final step of the proposed segmentation algorithm, the obtained probability map for vessel region is converted to a binary segmentation mask. The binarization is done by applying a threshold on the values in final probability map, where pixels with the probabilities higher than the threshold are labeled as vessel region.

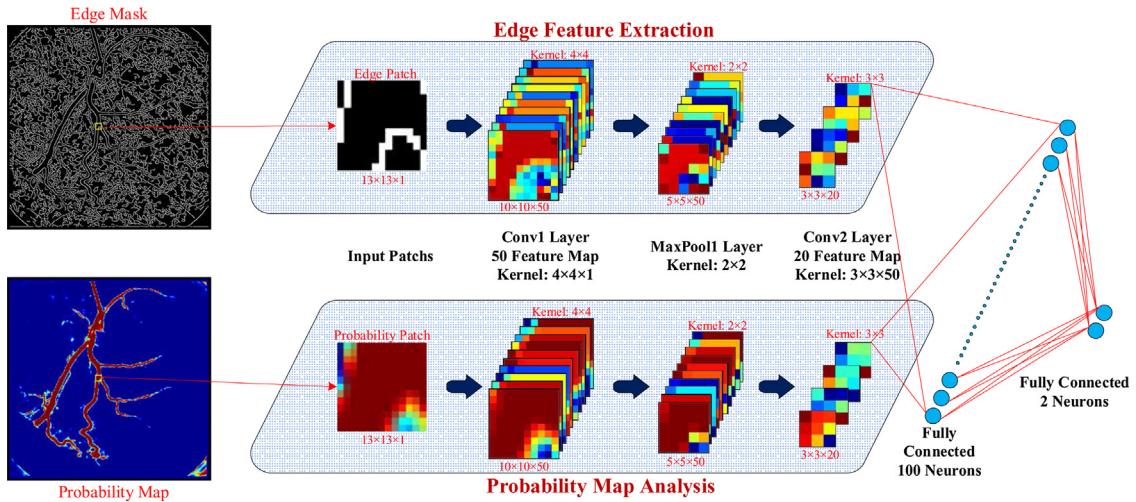


Fig. 5. Architecture of the second CNN stage.

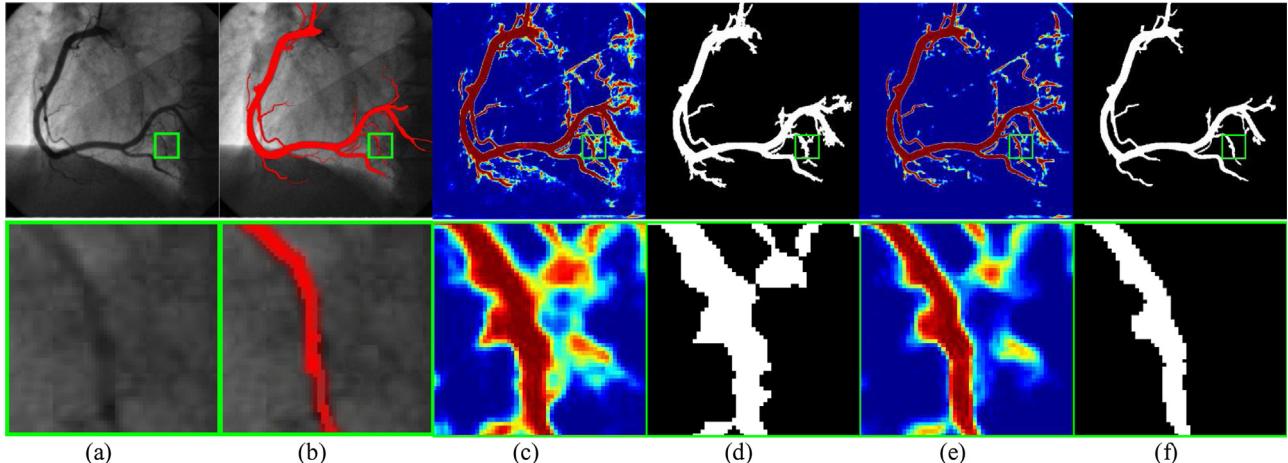


Fig. 6. Comparison of results of the first and the second stage of the proposed method. (a) input image, (b) ground truth segmentation, (c) CNN1 probability map, (d) CNN1 segmentation, (e) CNN2 probability map and (f) CNN2 segmentation.

In the proposed method an adaptive approach is used to find a proper threshold for each image.

For this aim, the threshold is initialized to a low value of $\varepsilon > 0$, e.g. $\varepsilon = 0.001$. In the initial binary mask, using this threshold, the biggest connected component is tracked. Afterward, the threshold is gradually increased and the resulted binary mask is obtained. By increasing the threshold, fewer points are selected as the vessel region. The growth in the value of the threshold is stopped when less than 80% of the points with $probability > 0.5$ are part of the largest connected component. The obtained value would be selected as the threshold for production of the final segmentation mask. This adaptive selection of the threshold is applied in order to maintain the connectivity of the extracted vessel regions.

3. Experimental results

The proposed method is evaluated on a dataset of X-ray angiography images. The dataset consists of 44 grayscale images sized at 512×512 pixels with available ground truth for segmentation by an expert. The proposed method is implemented in Matlab and Caffe [27]. The software runs on a computer system with Intel Core i7-4790K processor, 32 GB of RAM, and NVIDIA GeForce GTX Titan X GPU card. The average runtime of the proposed method for an image of size 512×512 in the applied dataset is 106 s.

The structuring element used for the multi scale Top-Hat transform in the preprocessing stage is a disk shape, which varies in size from 3 to 19. The solver type of stochastic gradient descent (SGD) is used for training and Xavier method [28] is used for weight initialization in the CNNs. Bias values are initially set to zero. The activation function of the fully connected layers is rectifier linear unit (ReLU). ReLU method results in faster convergence of the SGD algorithm in comparison to the other types of activation functions such as *sigmoid* or *tanh* [29]. The images of the dataset are randomly split into four equal sized folds. Images of two folds are used for training and another set is used for validation. The test is performed on the unseen images in the remaining fold. This means that 50% of the images (22 images) are used for training, 25% for the validation (11 images), and the remaining 25% of images (11 images) are left for the test.

The early stop strategy based on the validation set is applied to determine when to stop the network's training. The experiment is repeated over these four folds. A set of 30,000 selected patches are extracted from each training image, therefore a total of 990,000 patches are used for the training of the CNNs in each run. To decide on a patch size for each stage of the CNNs we performed a number of experiments and came up with suitable sizes. To control the network's learning and avoid overfitting the dropout technique is used which is a common method in deep networks training [30].

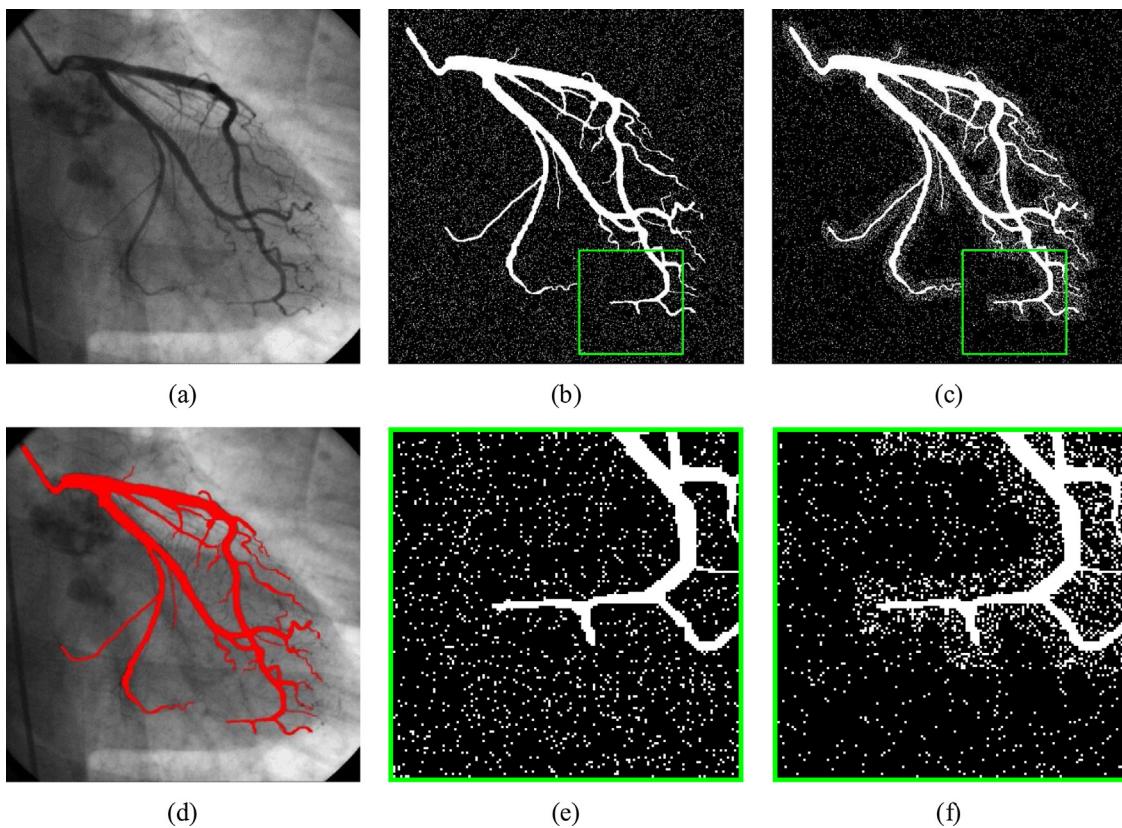


Fig. 7. Training patches of the CNN, (a) input image, (b, e) center pixels of the selected training patches are shown with white dots for CNN1 and (c, f) for CNN2, (d) ground truth.

The ratio of the dropout layer is set to 0.5, which means half of the neurons in the first fully connected layer are randomly chosen to be ignored in each training-iteration. This method would avoid the risk of over noticing a specific set of features in the training procedure.

3.1. Training patches

In angiography images, the vessel regions are a minor part of the image as compared to the image background. On average only about 4% of the pixels in these images are of the vessel regions. This means that by pure random selection of training patches, a majority of the selected patches (about 96%) would be such that the center pixels of the patches would belong to the background. This unbalanced distribution of training patches would reduce the learning performance of the network.

In the proposed method a policy for the selection of training patches is applied in order to avoid the problem of having unbalanced training data. In the selection of training samples for CNN1, 50% of patches are randomly selected from vessel region and the other 50% are extracted from the background. The selection is done based on the available ground truth of the training images. A similar procedure is used for CNN2, where 50% of patches are randomly selected from the vessel parts, and 30% from the areas around the vessels in the background. The other 20% are randomly selected from other parts of the background. The selection of patches in the areas around the vessel is with the aim of enhancing the network's accuracy in edge regions. Center pixels of the selected training patches in a sample training image for CNN1 and CNN2 are shown in Fig. 7. A total of 30,000 patches are extracted from the training image, where 15,000 are in the vessel regions and the other 15,000

patches are from the background. As can be seen, patches selected from the vessel parts have completely covered the vessel region.

3.2. Learned kernels

The learned filters in CNN1 are presented in Fig. 8. As mentioned before, CNN1 consists of two parallel paths that process local and global patches. The learned kernels in the first convolutional layer in these two paths are shown in Fig. 8. As can be seen, it seems that the global kernels have converged to track the general structure of the vessels, meanwhile the local kernels are learned to detect local edges in different directions.

The input patches are convolved with the learned kernels and the 50 feature maps are produced. Sample feature maps for three different regions as background, vessel parts, and vessel bifurcations, are shown in Fig. 9. As can be seen in Fig. 9(a), the existence of a vessel in a patch can be revealed by the global feature maps. Also, Fig. 9(c) reveals that discriminative features for exaction of locations of the vessel edges are incorporated in local feature maps. On the other hand, Fig. 9(b) shows that in the background patches, where no vessel part exists, the feature maps have a noisy characteristic.

3.3. Segmentation results

The proposed method is compared with state-of-the-art algorithms for extraction of vessel in X-ray angiograms [4,11,12]. Method of Ref. [4] is based on active contours, Ref. [11] is a filtering based model, and Ref. [12] uses graph cut for segmentation. Sample visual results of the methods are compared in Fig. 10 which also shows the existing ground truth for the segmentation. Presence of artifacts and variations of illuminance across the image back-

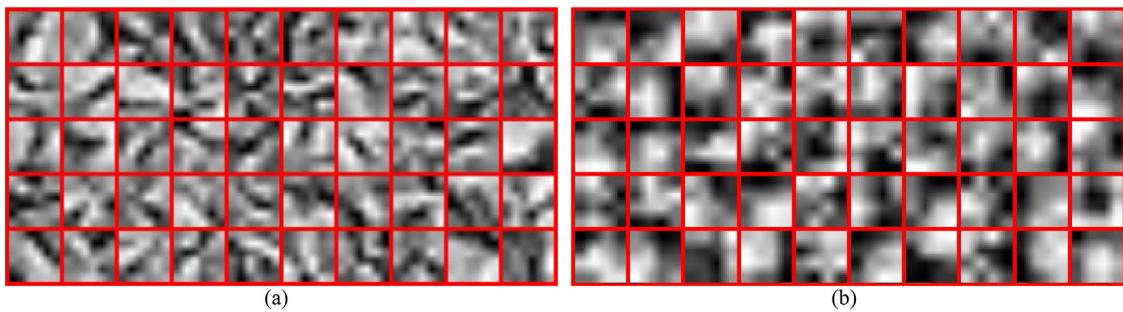


Fig. 8. Learned filters of the first convolutional layer in (a) global and (b) local path of CNN1.

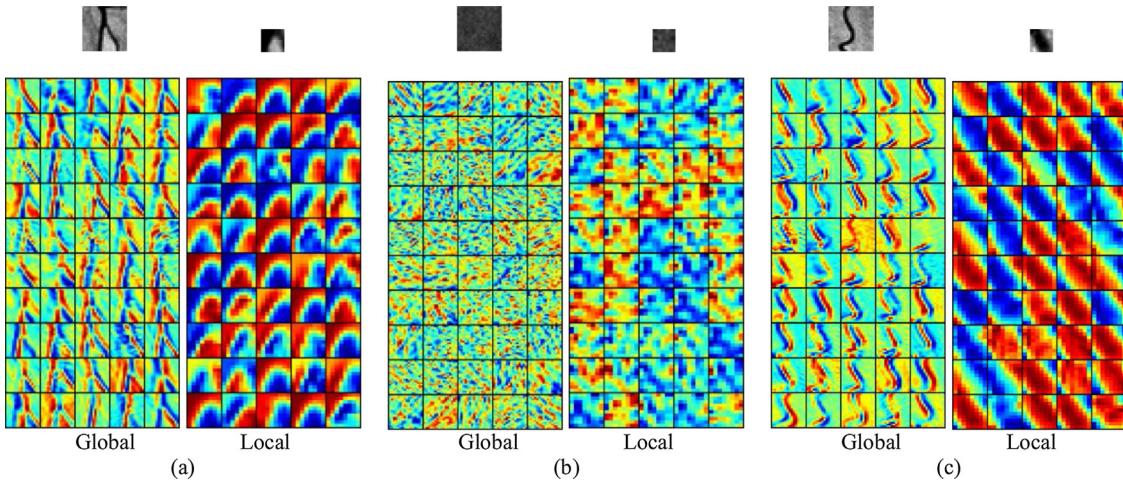


Fig. 9. Sample input patches and their corresponding feature maps. (a) Patches on bifurcations, (b) background patches and (c) vessel patches.

ground and foreground make the accurate extraction of vessels a challenging task.

In some images, such as the 4th column of Fig. 10, the other methods have missed some major arteries of the vessel. In the 1st and 5th columns of Fig. 10 other methods have wrongly labeled some artifacts as vessel regions and have failed accurate detection of the vessel edges. Meanwhile the proposed method produced promising results even in these complex cases.

For quantitative analysis, the segmentation performances of the proposed method is compared to the algorithms of Refs. [4,11,12,22] using metrics defined by Eqs. (9)–(13):

$$\text{accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (9)$$

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (11)$$

$$\text{dice_score} = \frac{2TP}{2TP + FN + FP} \quad (12)$$

$$\text{border error} = \frac{FP + FN}{TP + FN} \quad (13)$$

where TP (true positive) is the number of pixels that are labeled as vessel in both ground truth mask for segmentation and the produced segmentation mask. FN (false negative) is the number of vessel pixels that are missed by our segmentation mask. FP (false positive) counts the pixels that are falsely classified as vessel and finally TN (true negative) refers to number of background pixels that are truly labeled as background by our mask. These regions are illustrated in Fig. 11 for a sample vessel part, showing the available ground truth and the produced segmentation mask.

The results of metrics indicated by Eqs. (9)–(13) are presented in Table 1. The work of Ref. [22] is based on deep learning and uses one CNN. In the current proposed method, which uses two CNNs, the performance of the first CNN is improved by simultaneous usage of local and global patches. Also, the produced probability map of the first CNN is enhanced by addition of a second CNN. The second CNN

Table 1

Numerical comparison of segmentation methods, best results are in bold.

Algorithm	Segmentation Performance					
	Dice Score	Accuracy	Sensitivity	Specificity	Border Error	
Formula	$\frac{2TP}{2TP + FN + FP}$	$\frac{TP + TN}{TP + FN + TN + FP}$	$\frac{TP}{TP + FN}$	$\frac{TN}{TN + FP}$	$\frac{FP + FN}{TP + FN}$	
[12]	69.75	97.20	63.51	99.03	53.99	
[4]	68.50	96.91	67.88	98.41	62.03	
[11]	72.79	97.09	74.92	98.32	56.18	
[22]	75.62	97.27	79.35	98.91	49.24	
Proposed Method	81.51	97.93	86.76	98.59	39.70	

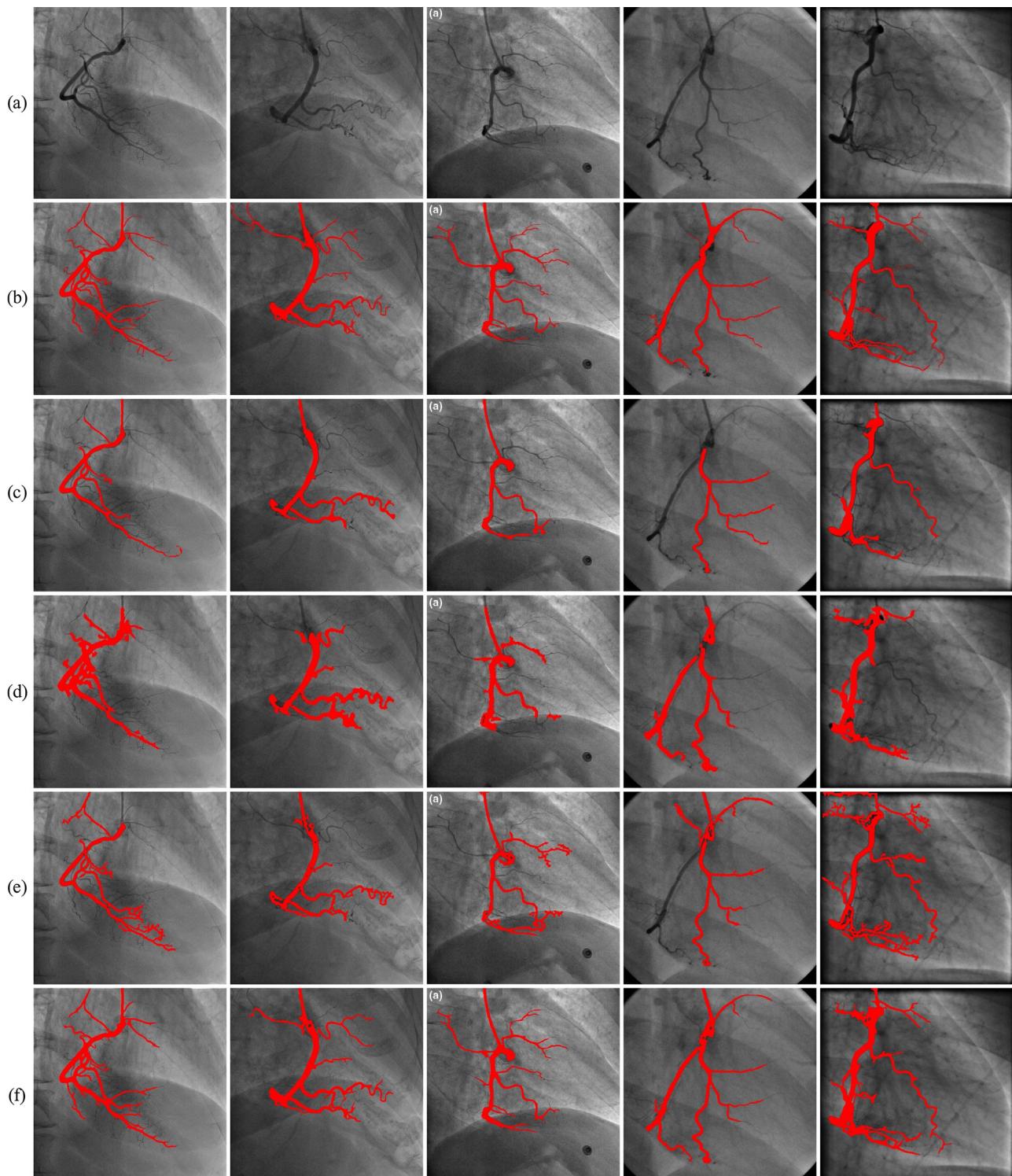


Fig. 10. Sample vessel extraction results, (a) original images, (b) ground truth, (c) results of Ref. [12], (d) results of Ref. [4], (e) results of Ref. [11] and (f) results of the proposed method.

uses the canny extracted edge patches and it takes the surrounding pixels' labels that are obtained from the first CNN's probability map. As can be seen in Table 1, the proposed method outperforms other algorithms in segmentation accuracy, sensitivity, border error and Dice score. It is noteworthy that due to precise detection of vessel parts, a high Dice score of 81.51% is achieved by the proposed method. This high Dice score surpasses the second best method by a difference of about 6%.

4. Conclusion

In this paper an angiogram vessel segmentation method was proposed. Angiograms lack uniformity of illumination and suffer from the presence of artifacts and noise. Hence, most conventional segmentation methods fail to perform accurately in such challenging situations. The proposed method used convolutional neural networks, as a deep learning tool, to produce an initial seg-

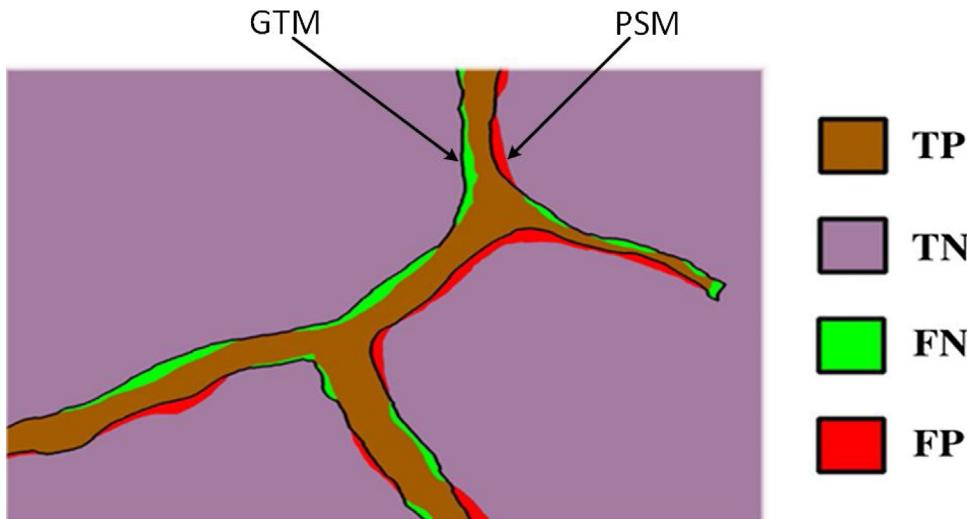


Fig. 11. Regions for measuring segmentation performance, where ground truth mask (GTM) is compared with the produced segmentation mask (PSM). TP and TN are truly classified pixels of vessels and background, FN and FP are falsely classified pixels of vessels and background respectively. Algorithm.

mentation probability map. The authors noticed that in the initial probability map there is a lack of accuracy in the border regions of the vessels. The accuracy of this initial map was improved by enhancing the contrast of the input image using a top-hat filter. For each pixel in the image a local patch was sent to the network to exploit local information about the pixel. The local information of the pixel helped accurate extraction of vessel boarders. Furthermore, for each pixel a global patch was extracted and fed to the network to help reveal the global structure of the vessels. Hence, the combined use of both local and global view of the image improved the segmentation probability map. The output probability map of the first CNN was further improved by using a second CNN. The second CNN was fed with the initial probability map and an edge map, which was extracted by canny edge-extraction routine. The second CNN knows the probabilities of the neighboring pixels from the initial map and also it knows the edge information of that pixel. Hence, it was possible for the second CNN to extract an accurate final probability map. The probability map was finally binarized to show the vessel region in the angiogram.

The accuracy and sensitivity of the proposed method were 97.93 and 86.76, respectively. The proposed method in terms of sensitivity was 7% better than the next best state of the art method. These innovations created a new system that outperformed the existing state-of-the-art segmentation methods by achieving almost 6% higher Dice-score.

Acknowledgment

The authors would like to thank Dr. Antonio Hernández Vela for sharing the source code of vessel segmentation [12], with us.

References

- [1] Coronary Artery Disease, 2016, <http://www.webmd.com/heart-disease/guide/heart-disease-coronary-artery-disease>. (Accessed 10 January 2017).
- [2] H. Fazlali, N. Karimi, S. Soroushmehr, S. Sinha, S. Samavi, B. Nallamothu, Vessel region detection in coronary X-ray angiograms, *IEEE Int. Conf. Image Process.* (2015) 1493–1497.
- [3] Y.C. Tsai, H.J. Lee, M.Y.C. Chen, Automatic segmentation of vessels from angiogram sequences using adaptive feature transformation, *Comput. Biol. Med.* 62 (2015) 239–253.
- [4] M.T. Dehkordi, A.M.D. Hoseini, S. Sadri, H. Soltanianzadeh, Local feature fitting active contour for segmenting vessels in angiograms, *IET Comput. Vis.* 8 (2014) 161–170.
- [5] Y. Yin, M. Adel, S. Bourennane, Retinal vessel segmentation using a probabilistic tracking method, *Pattern Recogn.* 45 (2012) 1235–1244.
- [6] Z. Shoujun, Y. Jian, W. Yongtian, C. Wufan, Automatic segmentation of coronary angiograms based on fuzzy inferring and probabilistic tracking, *Biomed. Eng. Online* 9 (2010).
- [7] F. Benmansour, L.D. Cohen, Tubular structure segmentation based on minimal path method and anisotropic enhancement, *Int. J. Comput. Vis.* 92 (2011) 192–210.
- [8] S. Cetin, G. Unal, A higher-order tensor vessel tractography for segmentation of vascular structures, *IEEE Trans. Med. Imaging* 34 (2015) 2172–2185.
- [9] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, Multiscale vessel enhancement filtering, *Int. Conf. Med. Image Comput. Assist. Interv.* (1998) 130–137.
- [10] M.A. Khan, M.K. Khan, M.A. Khan, Coronary angiogram image enhancement using decimation-free directional filter banks, *IEEE Int. Conf. Acoust. Speech Signal Process.* 5 (2004) 441–444.
- [11] B. Felfelian, H. Fazlali, N. Karimi, S. Soroushmehr, S. Samavi, et al., Vessel segmentation in low contrast x-ray angiogram images, *IEEE Int. Conf. Image Process.* 37 (2016) 375–379.
- [12] A. Hernández-Vela, C. Gatta, S. Escalera, L. Igual, V. Martin-Yuste, M. Sabate, et al., Accurate coronary centerline extraction caliber estimation, and catheter detection in angiographies, *IEEE Trans. Inf. Technol. Biomed.* 16 (2012) 1332–1340.
- [13] Y.Y. Boykov, M.-P. Jolly, Interactive graph cuts for optimal boundary & region segmentation of objects in ND images, *IEEE Int. Conf. Comput. Vis.* 10 (2001) 105–112.
- [14] M. Meliňščak, P. Prentašić, S. Lončarić, Retinal vessel segmentation using deep neural networks, *Int. Conf. Comput. Vis. Theory Applic.* 57 (2015) 577–582.
- [15] K.L. Hua, C.H. Hsu, S.C. Hidayati, W.H. Cheng, Y.J. Chen, Computer-aided classification of lung nodules on computed tomography images via deep learning technique, *OncoTargets Ther.* 8 (2015) 2015–2022.
- [16] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans. Med. Imaging* 35 (2016) 1207–1216.
- [17] M. Hawaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, et al., Brain tumor segmentation with deep neural networks, *Med. Image Anal.* 35 (2016) 18–31.
- [18] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. Soroushmehr, M. Jafari, K. Ward, et al., Melanoma detection by analysis of clinical images using convolutional neural network, *IEEE Int. Conf. Eng. Med. Biol. Soc.* 137 (2016) 1373–1376.
- [19] M. Jafari, N. Karimi, E. Nasr-Esfahani, S. Samavi, S. Soroushmehr, K. Ward, K. Najarian, Skin lesion segmentation in clinical images using deep learning, *Int. Conf. Pattern Recogn.* 33 (2016) 337–342.
- [20] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* 86 (1998) 2278–2324.
- [21] Michael A. Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015.
- [22] E. Nasr-Esfahani, S. Samavi, N. Karimi, S. Soroushmehr, K. Ward, M. Jafari, et al., Vessel extraction in X-Ray angiograms using deep learning, *IEEE Int. Conf. Eng. Med. Biol. Soc.* 64 (2016) 643–646.
- [23] X. Bai, F. Zhou, B. Xue, Image enhancement using multi scale image features extracted by top-hat transform, *Opt. Laser Technol.* 44 (2012) 328–336.
- [24] X. Bai, Morphological image fusion using the extracted image regions and details based on multi scale top-hat transform and toggle contrast operator, *Digital Signal Process.* 23 (2013) 542–554.

- [25] Y. Li, B. Yong, H. Wu, R. An, H. Xu, J. Xu, et al., Filtering airborne lidar data by modified white top-hat transform with directional edge constraints, *Photogramm. Eng. Remote Sens.* 80 (2014) 133–141.
- [26] M.A. Oskoei, H. Huosheng, A Survey on Edge Detection Methods, Technical Report: CES-506, School of Computer Science & Electronic Engineering, University of Essex, U.K, 2010.
- [27] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., Caffe: convolutional architecture for fast feature embedding, *ACM Int. Conf. Multimed.* 67 (2014) 675–678.
- [28] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feed forward neural networks, *Int. Conf. Artif. Intell. Stat.* 24 (2010) 249–256.
- [29] A. Krizhevsky, I. Sutskever, G. Hinton, Imagenet classification with deep convolutional neural networks, *Adv. Neural Inf. Process.* 109 (2012) 1097–1105.
- [30] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* 15 (2014) 1929–1958.