Vessel Detection in Ultrasound Images Using Deep Convolutional Neural Networks

2 authors, including:

Erik Smistad
SINTEF
51 PUBLICATIONS   **958** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   FAST - An open-source framework for high performance medical image computing and visualization   View project

Vessel detection in ultrasound images using deep convolutional neural networks [*]

Erik Smistad[1,2] and Lasse Løvstakken[1]

[1] Norwegian University of Science and Technology, Trondheim, Norway
[2] SINTEF Medical Technology, Trondheim, Norway
{erik.smistad, lasse.lovstakken}@ntnu.no

Abstract. Deep convolutional neural networks have achieved great results on image classification problems. In this paper, a new method using a deep convolutional neural network for detecting blood vessels in B-mode ultrasound images is presented. Automatic blood vessel detection may be useful in medical applications such as deep venous thrombosis detection, anesthesia guidance and catheter placement. The proposed method is able to determine the position and size of the vessels in images in real-time. 12,804 subimages of the femoral region from 15 subjects were manually labeled. Leave-one-subject-out cross validation was used giving an average accuracy of 94.5%, a major improvement from previous methods which had an accuracy of 84% on the same dataset. The method was also validated on a dataset of the carotid artery to show that the method can generalize to blood vessels on other regions of the body. The accuracy on this dataset was 96%.

1 Introduction

Blood vessel segmentation in ultrasound images may be useful in medical applications such as deep venous thrombosis detection [5], anesthesia guidance [12] and catheter placement. The goal of vessel detection in this work, was to identify the position and size of blood vessels in the image. Several segmentation and tracking methods require this as an initialization [1, 6]. In [12], a real-time vessel detection method was introduced, removing the need for manual initialization. This method performs an ellipse fitting at each pixel in the image using a graphic processing unit (GPU). However, this method has problems distinguishing vessels from non-vessels when varying user settings, such as gain, on the ultrasound scanner, and on individuals with more subcutaneous fat tissue, due to increased amounts of reverberation artifacts. Also, this method was only made to detect a single vessel for each image.

In this paper, we propose to use a similar ellipse fitting method to find vessel candidate regions which are passed on to a deep neural network classifier which determines if the region contains a vessel or not. As the proposed detection method provides both position and size, it may also be used as a vessel

---

segmentation method, assuming the vessel has an elliptical shape. The proposed method also enables detection of multiple vessels at the same time.

2 Methods

The next section will introduce the elliptic vessel model, which was used to find vessel candidate regions in the ultrasound image. Subimages were created from the ultrasound image for each vessel candidate. A deep convolutional network determines if each subimage is of an actual blood vessel. Fig. 1 provides an overview of the steps involved in the proposed method.
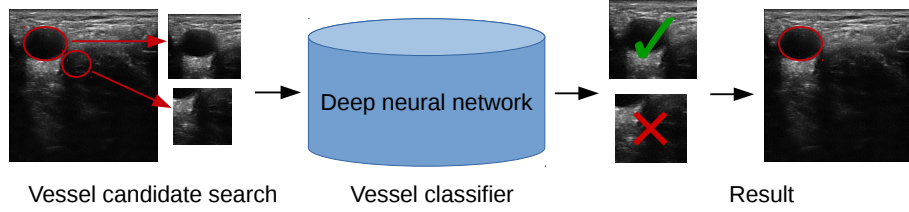

Vessel candidate search   Vessel classifier   Result

Fig. 1: Overview of the proposed method. The first step finds vessel candidates and creates subimages for each. The subimages are then passed on to a deep neural network which identifies the subimages belonging to vessels, and discards those that are not of vessels.

2.1 Vessel model

Each vessel is modelled as an ellipse with center $\boldsymbol{c} = [c_x, c_y]$ and major and minor radius a and b. The point $\boldsymbol{p}_i$ and its normal $\boldsymbol{n}_i$ of point i on an ellipse of N evenly distributed points can be calculated with the following equations.

$$\alpha_i = \frac{2\pi i}{N} \tag{1}$$

$$\boldsymbol{d}_i = [a\cos(\alpha_i), b\sin(\alpha_i)] \tag{2}$$

$$\boldsymbol{p}_i = \boldsymbol{c} + \boldsymbol{d}_i \tag{3}$$

$$\boldsymbol{n}_i = \frac{[b\cos(\alpha_i), a\sin(\alpha_i)]}{|[b\cos(\alpha_i), a\sin(\alpha_i)]|} \tag{4}$$

2.2 Vessel candidate search

First, the image is blurred using convolution with a Gaussian mask ($\sigma = 0.5mm$) and then the image gradients $\boldsymbol{G}$ are calculated using a central difference scheme.

For a given radii a and b, the vessel score S is calculated as the average dot product of the outward normal $\boldsymbol{n}_i$ and the corresponding image gradient at N points on the ellipse, as shown in (5).

$$S(\boldsymbol{c}, a, b) = \frac{1}{N} \sum_{i=0}^{N-1} \boldsymbol{n}_i \cdot \boldsymbol{G}(\boldsymbol{p}_i) \tag{5}$$

For each pixel, ellipses of different major radius a ranging from 3.5 to 6 mm, flattening factor f from 0 to 0.5 (minor radius $b = (1-f)a$) and $N = 32$ samples were used to calculate the vessel score. An increment of 0.25 mm was used for the radius, and 0.1 for the flattening factor. The ellipse with the highest score is selected for each pixel. The best score and the values a and b is stored for each pixel. Any vessel candidate with a score below 1.5 is discarded. This is a low threshold, which will not discard vessels with low contrast, but will also allow several non-vessel regions. Next, the vessel candidates are sorted according to their score from high to low. These are then processed in order, and a vessel candidate is accepted if the center is not inside another vessel candidate structure already accepted. Any vessel candidates which overlap with previously accepted vessel candidates are discarded.

For each vessel candidate, a square subimage is created from the ultrasound images as shown in Fig. 1. Examples of vessel candidates images are shown in Fig. 2. The vessel candidate image is centered at the vessel center $\boldsymbol{c}$ and the size of the image is determined by the major radius a so that the width and height of the image is $4a \times 4a$ converted to pixels. This image size will thus include the vessel as well as some surrounding tissue.

2.3   Vessel classifier

The next step of the proposed method is to send each vessel candidate image through a deep convolutional neural network classifier to determine if the image belongs to a blood vessel. Caffe [7] was used as the underlying framework both for training and testing of the classifier, while the vessel candidate search was implemented with the FAST medical image computing framework [11].

Data The data used for training and validation was acquired by first scanning the femoral region of both legs of 15 subjects with varying image quality and different ultrasound acquisition settings. Every tenth frame was run through the vessel candidate search step and the resulting images were stored on disk. This resulted in 12,804 images in total. All images were resized to 128×128, and classified manually as either vessel or non-vessel. Fig. 2 show some image examples of both blood vessels and non-vessel structures. The ultrasound system used was an Ultrasonix SonixMDP (Analogic, Boston, USA) with L14-5 linear array probe. To increase the amount of training data, all vessel candidate images were flipped horizontally, effectively doubling the amount of training data.
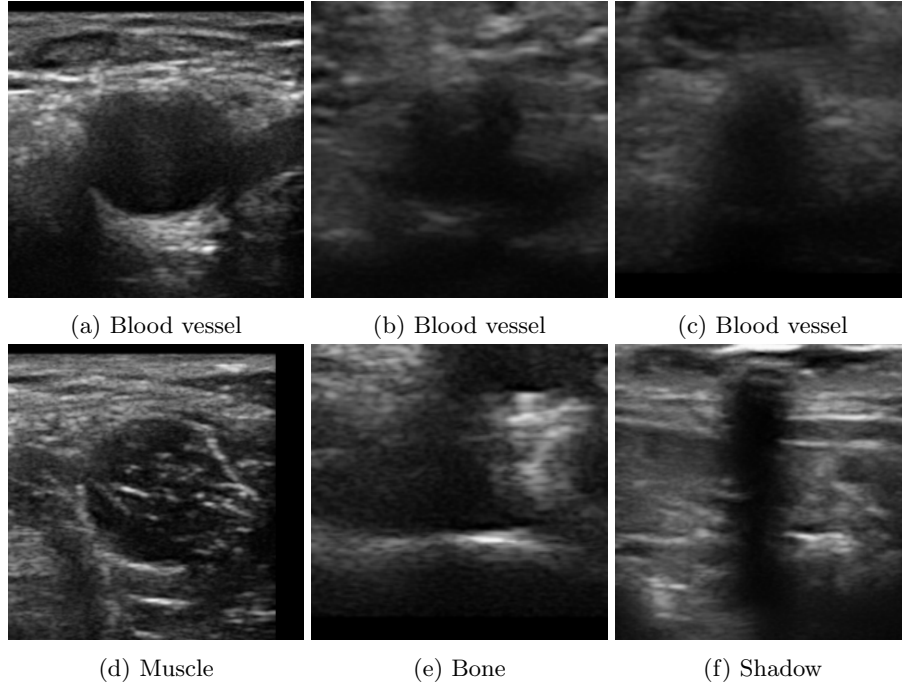
Fig. 2: Vessel candidate images of blood vessels and other non-vessel structures used for training the neural network.

The network The AlexNet [8] network was used initially, and gradually simplified by removing convolution-pooling blocks and reducing the number of convolutions, while maintaining the validation accuracy. The network was simplified mainly to improve the test runtime speed, which was important in order to achieve real-time performance. The final vessel classification network consisted of two convolution layers, one normalization layer, two max pooling layers and three fully connected layers. Additionally, rectified linear units (ReLU), which have shown to improve training [4], was used as non-linear activation units both for the convolution layers and the fully connected (FC) layers. Thus, including ReLU layers, the network consisted of 13 layers in total, as shown in Fig. 3. Additionally, a softmax loss layer was used for the training of the network. The data layer size was fixed to 110×110 pixels. During training, random patches of size 110×110 were cropped from the 128×128 vessel candidate images to prevent overfitting. This technique increased accuracy with about 1%. The mean image was calculated from the training data and subtracted from the input image. The first convolution layer had 9 convolutions of size 11×11 pixels and the second had 32 convolutions of size 15×15. The max pooling was done over patches of 3×3. Local response normalization (LRN) [8] was used after the first convolution layer with the same parameters as in [8]. Dropout was used on the

fully connected layers with a probability of 0.5. The network was trained with stochastic gradient descent, batch size 128, momentum 0.9 and weight decay 0.0005. The base learning rate was 0.01 with a sigmoid learning rate decay.
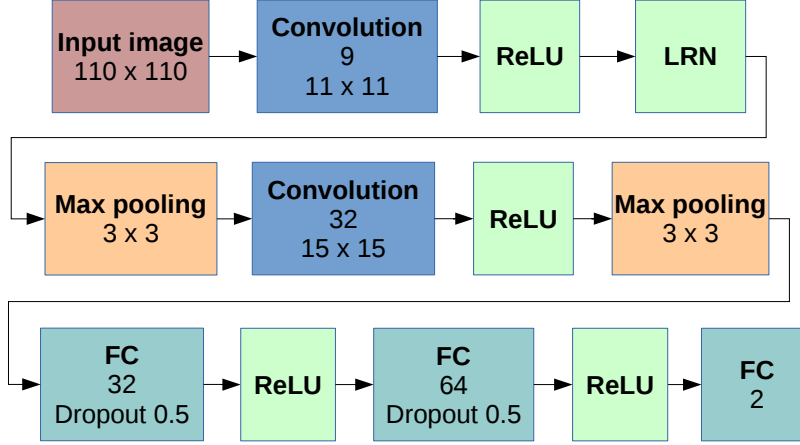
Fig. 3: The vessel detection network. A fixed-size input image of size 110×110 is feed into two convolution-pooling stages with 9 and 32 convolutions respectively. This is followed by three fully connected (FC) layers with dropout to reduce overfitting. A local response normalization (LRN) is performed after the first convolution. Rectified linear units (ReLU) are used as non-linear activation units in all stages.

2.4 Performance optimizations

Ultrasound is a real-time imaging modality, delivering typically 10-20 images per second. The proposed method thus have to find vessels in each image in less than 100 ms to be able to process the ultrasound image stream in real-time. The vessel candidate search, subimage creation and resizing were all implemented on the GPU using the FAST framework. Caffe was run in GPU mode and all vessel candidates for a given image frame were batch processed, which significantly boosts performance. Additionally, the vessel candidate search was only performed on every fourth pixel.

3 Results

Fig. 4 show the convolutions learned by the neural network. These figures show that the first convolutional layer learns to detect horizontal edges, and the second layer learns to identify different patterns of horizontal edges. The trained neural

network does not seem to find vertical edges as important in the ultrasound images. This seems sensible, as vertical edges are often weaker or missing in ultrasound images.

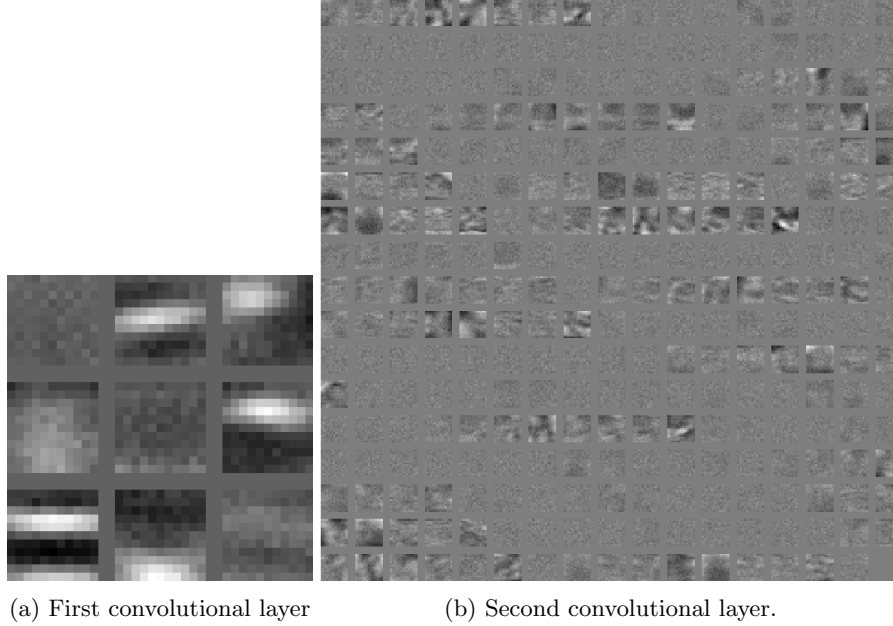(a) First convolutional layer   (b) Second convolutional layer.

Fig. 4: Features learned by the neural network. The first layer has learned several horizontal edge detectors, while the second convolutional layer has learned to recognize patterns of horizontal edges.

Leave-one-subject-out cross validation was used, thus 14 subjects were used for training and 1 subject kept for validation. The average classification accuracy for the cross validation was 94.5%, with a standard deviation of 2.9. This was calculated using a discrimination threshold of 0.5 on the softmax output of the vessel classifier. Fig. 5a shows the result of the vessel detection on an image of the femoral region.

This dataset was only from a single area of the body, the femoral region covering the femoral artery and vein. To see how well the proposed method can generalize to other parts of the body, a dataset of the left and right carotid artery was acquired from two subjects and used as validation data, while the dataset with the 15 subjects of the femoral region was used as training data. The dataset was created with the same method described in section 2.3. With this data, the method achieved an accuracy of 96%. Fig. 5b shows the result of the vessel detection on an image of the carotid artery.
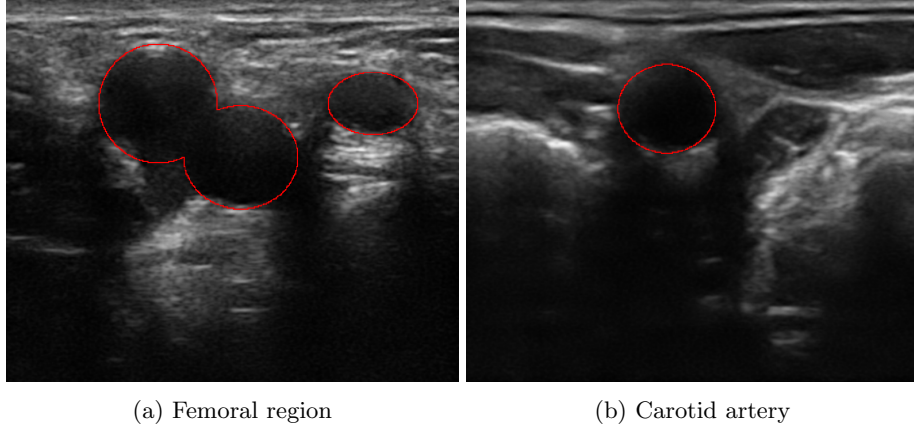
(a) Femoral region (b) Carotid artery

Fig. 5: Vessel detection result on two ultrasound images.

The proposed method was compared to a another state of the art vessel detection method [12]. This method achieved an average accuracy of 84% on the femoral region dataset. The receiver operating characteristics (ROC) curves in Fig. 6 show how the two methods perform when varying the discrimination threshold for the same dataset.
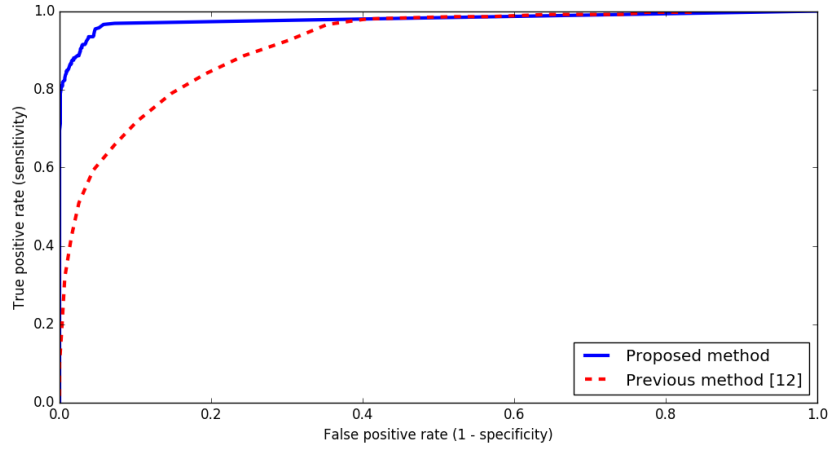

Fig. 6: ROC curve of the proposed method and the method in [12].

Training time was about 10 minutes on a laptop computer with an NVIDIA GTX 980M GPU with 8 GB of memory. The average runtime of all steps including the vessel candidate search and vessel classification was 46 ms, enabling the ultrasound images to be processed in real-time.

4 Discussion

The vessel model used in the proposed method assumes that the vessels are elliptical, while this often holds true for arteries, it may not be ideal for veins which often have a more irregular shape. Thus, the proposed method is more suited for arteries than veins. The vessel model also does not consider rotation of the ellipse. However, in our experiments this has not been an issue as vessels usually are compressed in the vertical direction, due to pressure from the ultrasound probe applied by the user. Including rotation in the vessel candidate search would significantly reduce runtime performance.

An alternative to the proposed ellipse fitting method would be to use more general object detection methods, such as R-CNN [3, 10]. However, these methods are more complex and bounding boxes would have to be created manually around each vessel in each image, which is time consuming. With the ellipse fitting method, the user only have to choose between the classes "vessel" and "non-vessel" for each vessel candidate subimage. Thus, the proposed ellipse fitting method aids in the labeling of the data.

Another alternative can be to use a fully convolutional neural network [9]. Such as network would provide a classification of each pixel. The ground truth data could be created by a user selecting the center of each blood vessel. Using such as network may be more robust in terms of rotation and deformation of the blood vessels. However, it would not provide the radius of the vessels and a segmentation as shown in Fig. 5.

The validation accuracy was 94.5%, which is a major improvement from the vessel detection method in [12] which got an accuracy of 84% on the same validation dataset. The accuracy may be improved by adding more training data, including the temporal dimension of the data with recurrent neural networks, and including Doppler data in a separate image channel. In the proposed network, the weights were initialization using Gaussian noise with standard deviation 0.01. Unsupervised pre-training has shown to be a good way to initialize the weights of deep neural networks [2]. With ultrasound imaging, a large amount of unlabeled data can easily be acquired from the target body regions. Thus, we believe unsupervised pre-training of deep networks will be a useful technique within ultrasound imaging.

5 Conclusion

A robust real-time vessel detection method for ultrasound images was presented. The method uses a deep convolutional neural network to classify subimages. Although the neural network was only trained on images of the femoral artery and vein, it is able to generalize to other vessels such as the carotid artery.

References

1. Abolmaesumi, P., Sirouspour, M., Salcudean, S.: Real-time extraction of carotid artery contours from ultrasound images. In: Proceedings 13th IEEE Symposium

on Computer-Based Medical Systems. CBMS 2000. pp. 181–186. IEEE Comput. Soc (2000)

2. Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H.: Greedy Layer-Wise Training of Deep Networks. Advances in Neural Information Processing Systems 19(1), 153–160 (2007)

3. Girshick, R.: Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision (ICCV). pp. 1440–1448. IEEE (dec 2015)

4. Glorot, X., Bordes, A., Bengio, Y.: Deep Sparse Rectifier Neural Networks. 14th International Conference on Artificial Intelligence and Statistics pp. 315–323 (2011)

5. Guerrero, J., Salcudean, S.E., McEwen, J.A., Masri, B.A., Nicolaou, S.: System for Deep Venous Thombosis Detection Using Objective Compression Measures. IEEE Transactions on Biomedical Engineering 53(5), 845–854 (2006)

6. Guerrero, J., Salcudean, S.E., McEwen, J.A., Masri, B.A., Nicolaou, S.: Real-time vessel segmentation and tracking for ultrasound imaging applications. IEEE Transactions on Medical Imaging 26(8), 1079–1090 (2007)

7. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional Architecture for Fast Feature Embedding. Proceedings of the ACM International Conference on Multimedia pp. 675–678 (2014)

8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems pp. 1097–1105 (2012)

9. Long, J., Shelhamer, E., Darrell, T.: Fully Convolutional Networks for Semantic Segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3431–3440 (2015)

10. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in neural information processing systems pp. 91–99 (jun 2015)

11. Smistad, E., Bozorgi, M., Lindseth, F.: FAST: framework for heterogeneous medical image computing and visualization. International Journal of Computer Assisted Radiology and Surgery 10(11), 1811–1822 (2015)

12. Smistad, E., Lindseth, F.: Real-Time Automatic Artery Segmentation, Reconstruction and Registration for Ultrasound-Guided Regional Anaesthesia of the Femoral Nerve. IEEE Transactions on Medical Imaging 35(3), 752–761 (mar 2016)