

# Retinal Vessel Segmentation from Simple to Difficult

Qing Liu<sup>1,2</sup>, Beiji Zou<sup>1\*</sup>, Jie Chen<sup>2</sup>, Zailiang Chen<sup>1</sup>, Chengzhang Zhu<sup>1</sup>,  
Kejuan Yue<sup>1</sup>, Guoying Zhao<sup>2</sup>

<sup>1</sup> School of information science and engineering, Central South University

<sup>2</sup> Center for Machine Vision and Signal Analysis, University of Oulu

**Abstract.** In this paper, we propose two vesselness maps and a simple to difficult learning framework for retinal vessel segmentation which is ground truth free. The first vesselness map is the multiscale centreline-boundary contrast map which is inspired by the appearance of vessels. The other is the difference of diffusion map which measures the difference of the diffused image and the original one. Meanwhile, two existing vesselness maps are generated. Totally, 4 vesselness maps are generated. In each vesselness map, pixels with large vesselness values are regarded as positive samples. Pixels around the positive samples with small vesselness values are regarded as negative samples. Then we learn a strong classifier for the retinal image based on other 3 vesselness maps to determine the pixels with mediocre values in single vesselness map. Finally, pixels with two classifier supports are labelled as vessel pixels. The experimental results on DRIVE and STARE show that our method outperforms the state-of-the-art unsupervised methods and achieves competitive performances to supervised methods.

## 1 Introduction

Retinal fundus images provide a window to inspect the fundus of the eye, and they are widely used for the diagnosis of various pathologies, such as age related macular degeneration and diabetic retinopathy, glaucoma etc. Manual analysis of the retinal images is time consuming and expensive for ophthalmologists. Moreover, it is impossible to quantify the structures accurately in the fundus. Therefore, the automation of the analysis becomes important.

As one of the basic procedures in automatic analysis, vessel segmentation is still a challenge task. On one hand, the width of the vessels has large variability. For example, vessels at the end of each branch are always only several pixel width or even only one pixel width in the images. On the other hand, due to low image quality, the retinal images are noised and in various brightness.

---

\* Corresponding author: bjzou@csu.edu.cn. Q. Liu was supported by the scholarship from China Scholarship Council. B. Zou and Z. Chen were supported by the NSF of China under Grant No.61573380 and No. 61440055. G. Zhao and J. Chen were supported by Academy of Finland, Tekes Fidipro Program and Infotech Oulu.

Recently, several methods have been proposed to segment the vessels from the retinal fundus images, which can be classified into two classes, i.e., unsupervised methods and supervised methods.

1) Unsupervised Methods: The unsupervised methods model for the shape prior, appearance prior or the profile prior of the vessels. For example, Zana F. et al [1] propose to use mathematical morphology operators with linear structuring elements to model the vessels' morphological properties, such as linearity, connectivity. Mendonca A. M. et al. [2] use a top-hat transform with circular structuring elements to enhance the vessels first, then reconstruction operators are used to segment the vessels. Inspired by the profile prior of the vessel, matched filters such as trainable COSFIRE filters [3], multiscale line filters [4], first-order of derivative of Gaussian [5] are proposed to model the profile of vessels. Such methods rely on thresholding filter responses to obtain the vessel points. The performances are sensitive to the threshold selection. Usually, a large threshold results in numerous missing detection of vessel pixels while a small threshold leads to lots of wrong detection of background pixels. Moreover, they are limited when the vessel is slightly different with the expected pattern.

2) Supervised Methods: The supervised methods relying on the ground truth learn feature aggregation strategies to detect the vessels. For example, Ricci E. et al. [6] learn a support vector machine (SVM) to detect the vessel points from line filter responses. Lupascu C.A. et al. [7] learn an Adaboost classifier which takes multiscale local intensity structure, spatial properties and geometry features as input while [8] learn a LogitBoost classifier. In [9], gray-level and moment invariants based features are extracted to learn a neural network scheme for vessel segmentation. In [10], a deep neural network is trained to learn a cross-modality data transform from retinal image to vessel map. Supervised methods are more invariable to vessel deformations and brightness since they combine different priors about the vessels. However, in such methods, an extra time-consuming off-line learning process is required and the ground truth are necessary. It is obvious that professional skills are required to the tedious pixel level ground truth labelling. Besides, such methods are dataset dependant, and their performances decrease remarkably when testing on other datasets.

To alleviate the extra learning process and avoid ground truth labelling, we present a ground truth free learning framework to segment the vessels. First, two weak vesselness maps are proposed. One is called centreline-boundary contrast, modelling for the appearances of the vessels. The other is called the difference of diffusion which use the difference between the diffused image and the original one to measure the vesselness. Meanwhile, other two existing vesselness maps, i.e., the response map of B-COSFIRE filter [3] and line detector response map [4] are also extracted. Those weak vesselness maps are used to generate training samples for strong models. Second, for the training sample set from one weak vesselness map of one retinal image, a strong classifier is learnt to detect the vessel pixels. Totally, four classifiers are learnt for each retinal image and pixels those win two or more than two votes are classified as vessel pixels finally.

In the remainder of this paper, we first introduce the proposed method in

Section 2 in detail. Then we report our experimental results and the comparisons to the state of the arts in Section 3. Finally, we conclude our paper in Section 4.

## 2 Proposed Approach

The simple to difficult learning framework includes two steps. One is weak vesselness maps generation and the other is strong classifiers learning. The final results are obtained by voting.

### 2.1 Weak Vesselness Maps Generation

We extract vesselness features from the green channel of the RGB retinal images since the contrast between vessels and background in green channel is higher than in the red and blue ones. We extract 4 weak vesselness maps to generate the training samples for strong classifier learning. They are the output of the multiscale line filters ( $f_L$ ) [4], the output of B-COSFIRE filters ( $f_B$ ) [3], the multiscale centreline-boundary contrast map ( $f_C$ ), the difference of the diffusion map ( $f_D$ ). Since the first two vesselness maps have been described in [4] and [3] respectively, in the following, we describe the latter two features in detail.

#### *Multiscale Centreline-Boundary Contrast*

According to the properties that the intensities of the vessels are lower than the background pixels and the vessels seem to be linear locally, we present a centreline-boundary contrast filter with line structure to enhance the vessels. Fig. 1a shows a basic centreline-surround contrast filter with size  $5 \times 7$ . Generally, we define a basic  $(2r_1 + 1) \times (2r_2 + 1)$  centreline-boundary contrast filter with direction  $0^\circ$  by:

$$G(x, y; r_1, r_2, 0^\circ) = \begin{cases} 1/(4r_2 + 2) & \text{if } x = 1 \text{ or } x = (2r_1 + 1) \\ -1/(2r_2 + 1) & \text{if } x = r_1 + 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Vessels distribute over the whole retinal image and converge into the optic

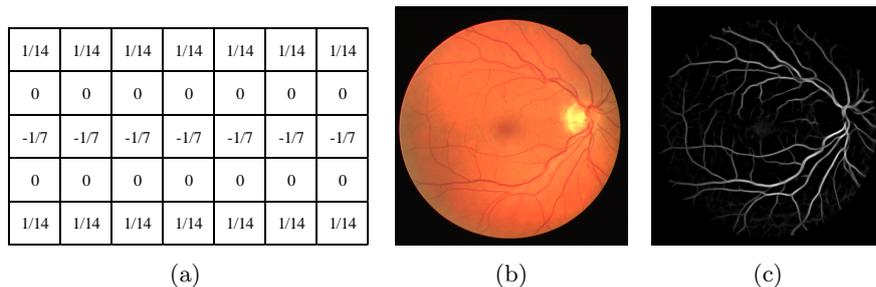


Fig. 1: (a) A  $5 \times 7$  centreline-boundary contrast filter with line structure. (b) The input image. (c) The centreline-boundary contrast vesselness map.

disk from different directions. Moreover, the width of vessels are variable. Usually, the width of the trunk vessels are larger than their branch vessels. To handle the vessels in different directions, we rotate  $G(\cdot; r_1, r_2, 0^\circ)$  every  $15^\circ$  from  $0^\circ$  to  $180^\circ$ , and generate 12 filters. To enhance the vessels with variable width, we further generalise the centreline-boundary contrast filter by varying the width  $r_1$  and  $r_2$  from 3 to 19 with step 2. Correspondingly, given an image  $I$ , the centreline-boundary contrast vesselness map is computed by:

$$f_C = \sum_{r_1, r_2} \max_{\theta} I \circ G(\cdot; r_1, r_2, \theta) \quad (2)$$

where  $\circ$  is a convolution operator. Thus, vessel-like pixels are enhanced due to their high contrast to background while the background pixels are suppressed for their smoothness. Fig. 1c shows an example of a multiscale centreline-boundary contrast map.

### ***Difference of Diffusion***

Regarding the image as a partial differential equation (PDE) based diffusion system, we mathematically formulate the image diffusion as an evolutionary PDE controlled by a function  $g(x, y, t)$ :

$$\frac{\partial I}{\partial t} = DIV(g(x, y, t) \nabla I) \quad (3)$$

where  $DIV$  is a divergence operator,  $g(x, y, t)$  is a non-linear diffusion function of  $t$ ,  $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$  is the spatial gradient. We adopt Tri-diagonal matrix algorithm to solve the diffusion equation [11]. The distribution of the intensities will reach equilibrium as time goes on. Since the background regions are almost homogeneous and they account for a large part of the image, the difference between their diffused state and their initial states are small. On the contrary, the vessel pixels always have lower intensities than their surrounding pixels and account for only a small part of the image. The difference between their diffused states and their initial states are remarkable. Therefore, we use the difference between the diffused image and the original one to measure the vesselness:

$$f_D = I_E - I \quad (4)$$

where  $I_E$  is the diffused image when  $t = 5$ .

## **2.2 Vessel Segmentation From Simple to Difficult**

Until now, we obtain 4 weak vesselness maps. We normalise each vesselness map into  $[0, 1]$ , and the vesselness value indicates the probability of the pixel to be a vessel pixel. In each vesselness map, the pixels with extremely large vesselness values can be easily detected as vessel pixels. Similarly, pixels with extremely small vesselness values can be easily detected as background pixels. We call those pixels are simple pixels. The labels of pixels with mediocre vesselness values are controversial only according to one single vesselness map. We call those pixels

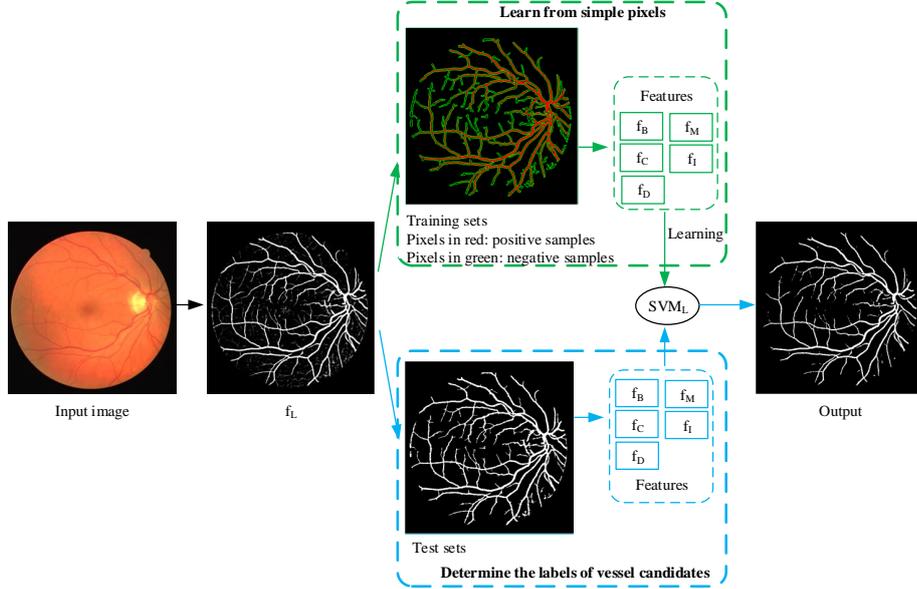


Fig. 2: Strong classifier learning on simple pixels and candidate pixels classification in  $f_L$ . The top  $T_1$  pixels in  $f_L$  consist of the positive samples. The top  $T_2$  pixels in  $f_L$  consist of the test set. The pixels around the test set consist of the negative samples. The learning process flow is marked in green arrows while the candidate pixels classification is marked in blue arrows.

difficult pixels. To determine the labels of the difficult pixels, we learn a strong classifier on simple pixels sampled from one single vesselness map leveraging to the other 3 vesselness maps.

Formally, for each weak vesselness map  $f_k \in \{f_L, f_B, f_C, f_D\}$ , we first decompose it into two sets: the vessel candidate set  $S_C$  and the background set  $S_B$ .  $S_C$  collects the top  $T_1$  pixels and  $S_B$  collects the rest pixels. Pixels in  $S_B$  are directly regarded as background pixels while the labels of pixels in  $S_C$  are given by a strong classifier. To learn the strong classifier, we first generate training samples from  $f_k$ . We take the top  $T_2$  pixels in  $f_k$  as positive samples, denoted as  $S_+$ . Pixels around the  $S_C$  are regarded as negative samples, denoted by  $S_-$ .  $S_+$  and  $S_-$  consist of the training data  $\{S_+ \cup S_-\}$ . We use the rest 3 vesselness maps  $\{f_L, f_B, f_C, f_D\}/f_k$  to describe the samples in the training set. Since the green intensity  $f_I$  itself and the local mean removed green intensity  $f_M$  also have discriminabilities, we add them into the feature vector. Thus, each sample is described by a 5-D feature vector. Given the training samples, we train a SVM classifier with radial basis function to determine whether the pixels in  $S_C$  are vessel points or not. Totally, we learn four classifiers and obtain four binary segmentations. To generate a final vessel segmentation, pixels that are supported by at least two classifiers are labelled as vessel pixels. Taking the weak vesselness map  $f_L$  as an example, Fig. 2 illustrates the proposed learning framework.

### 3 Experimental Results

We use two widely used public datasets to evaluate the proposed method. One is DRIVE [12] which includes 20 retinal images for training and 20 images for testing. For each image, two manually segmented maps by two observers and a mask image are provided. The other is STARE [13], which include 20 images and two corresponding manually segmentations by two different observers. In STARE, there are 10 images containing signs of pathologies and 10 healthy retinal images. In our experiments, we set  $T_1 = 12\%$ ,  $T_2 = 7\%$  for DRIVE and  $T_1 = 12\%$ ,  $T_2 = 5\%$  for STARE.

Following previous methods [3] [8] [10], we adopt accuracy (Acc), sensitivity (Se), specificity (Sp) to compare the performance of the proposed method with the state-of-the-art methods. The accuracy measures the proportion of pixels that are correctly detected. The sensitivity measure the proportion of vessel pixels that are correctly detected. The specificity is the proportion of background pixels that are correctly detected. Conventionally, only the pixels in the field of the view are taken into consideration since pixels out the field of the view can be easily labelled as background via thresholding.

To demonstrate the effectiveness of the proposed simple to difficult learning framework, we compare the final performances with the performances of each single component. Table. 1 shows that after our simple to difficult learning scheme, the sensitivities, specificities and accuracies on both two datasets are superior to each single component. Meanwhile, we can see that the two proposed vesselness maps, especially the centreline-boundary contrast map, also have strong discriminant ability.

We further compare our method with 4 supervised methods and 4 unsupervised methods. The results of the proposed method and the compared methods on DRIVE[12] and STARE[13] are reported in Table. 2 and Table. 3 respectively. Generally, the proposed method outperforms the unsupervised methods in terms of sensitivity, specificity and accuracy on both two datasets. Compared to supervised methods, the proposed method still achieves better sensitivity. The performance of the supervised methods [8] and [10] significantly outperforms the proposed method on DRIVE when the models are trained on the training images from DRIVE. However when the models are trained on STARE, their accuracies decrease to 0.9456 and 0.9486 respectively while ours is 0.9451. It's worth noting that the proposed method outperforms most of the supervised methods on STARE dataset except for [10]. Even though, our method achieves competitive

Table 1: Component analysis on DRIVE and STARE datasets

feature type	DRIVE			STARE		
	Se	Sp	Acc	Se	Sp	Acc
centreline-boundary contrast	0.7659	0.9604	0.9354	0.7037	0.9766	0.9493
difference of diffusion	0.7388	0.9561	0.9283	0.6217	0.9843	0.9470
line detector [4]	N.A	N.A	0.9324	N.A	N.A	0.9324
B-COSFIRE [3]	0.7655	0.9704	0.9442	0.7716	0.9563	0.9497
<b>proposed method</b>	<b>0.7718</b>	<b>0.9707</b>	<b>0.9451</b>	<b>0.7822</b>	<b>0.9745</b>	<b>0.9541</b>

accuracy to [10] when [10] trained the model on DRIVE. The visual comparisons between [3], [4] and our method are provided in our supplementary material.

## 4 Conclusion

In this paper, we present two features and a new framework for vessel segmentation in retinal images. Our framework learns classifiers from the pixels whose labels can be determined easily to determine the labels of the pixels that are controversial in a single feature map. Our method shows its advantages in two aspects: (1) comparing to unsupervised methods, our method avoids the sensitive threshold selection; (2) comparing to supervised methods, our method does not rely on the expensive ground truth. Our experimental results on two public datasets illustrate that the proposed method outperforms the state of the art unsupervised methods on DRIVE and STARE datasets. Compared to supervised methods, the proposed ground truth free method also achieves competitive performances.

Table 2: Comparisons with state-of-the-art methods on the DRIVE dataset.

type	methods	Se	Sp	Acc
	Second observer	0.7796	0.9717	0.9470
Supervised methods	Staal et al. (2004) [12]	N.A	N.A	0.9441
	Soares et al. (2006) [14]	0.7332	0.9782	0.9461
	Fraz et al. (2012) [8] trained on DRIVE	0.7406	0.9807	0.9480
	Fraz et al. (2012) [8] trained on STARE	0.7242	0.9792	0.9456
	Li et al. (2016) [10] trained on DRIVE	0.7569	0.9816	0.9527
	Li et al. (2016) [10] trained on STARE	0.7273	0.9810	0.9486
Unsupervised methods	Al-Diri et al. (2009) [15]	0.7282	0.9551	N.A
	Fraz et al. (2012) [16]	0.7152	0.9759	0.9430
	Nguyen et al. (2013) [4]	N.A	N.A	0.9407
	Azzopardi et al. (2015) [3]	0.7655	0.9704	0.9442
	<b>Proposed method</b>	<b>0.7718</b>	<b>0.9707</b>	<b>0.9451</b>

Table 3: Comparisons with state-of-the-art methods on the STARE dataset.

type	methods	Se	Sp	Acc
	Second observer	0.8951	0.9384	0.9348
Supervised methods	Staal et al. (2004) [12]	N.A	N.A	0.9516
	Soares et al. (2006) [14]	0.7207	0.9747	0.9479
	Fraz et al. (2012) [8] trained on STARE	0.7548	0.9763	0.9534
	Fraz et al. (2012) [8] trained on DRIVE	0.7010	0.9770	0.9495
	Li et al. (2016) [10] trained on STARE	0.7726	0.9844	0.9628
	Li et al. (2016) [10] trained on DRIVE	0.7027	0.9828	0.9545
Unsupervised methods	Al-Diri et al. (2009) [15]	0.7521	0.9681	N.A
	Fraz et al. (2012) [16]	0.7311	0.968	0.9442
	Nguyen et al. (2013) [4]	N.A	N.A	0.9324
	Azzopardi et al. (2015) [3]	0.7716	0.9563	0.9497
	<b>Proposed method</b>	<b>0.7822</b>	<b>0.9745</b>	<b>0.9541</b>

## References

1. Zana F, Klein JC. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE transactions on image processing* 10(7):1010-9 (2001)
2. Mendonca, A.M. and Campilho, A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE transactions on medical imaging*, 25(9):1200-1213 (2006)
3. Azzopardi, G., Strisciuglio, N., Vento, M. and Petkov, N. Trainable COSFIRE filters for vessel delineation with application to retinal images. *Medical image analysis*, 19(1) : 46-57 (2015).
4. Nguyen, U.T., Bhuiyan, A., Park, L.A. and Ramamohanarao, K. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern recognition*, 46(3) : 703-715 (2013).
5. Zhang, B., Zhang, L., Zhang, L. and Karray, F. Retinal vessel extraction by matched filter with first-order derivative of Gaussian. *Computers in biology and medicine*, 40(4) : 438-445 (2010).
6. Ricci, E. and Perfetti, R. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE transactions on medical imaging*, 26(10) : 1357-1365 (2007).
7. Lupascu, C.A., Tegolo, D. and Trucco, E. FABC: retinal vessel segmentation using AdaBoost. *IEEE Transactions on Information Technology in Biomedicine*, 14(5) : 1267-1274 (2010).
8. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.R., Owen, C.G. and Barman, S.A. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9) : 2538-2548 (2012).
9. Marn, D., Aquino, A., Gegndez-Arias, M.E. and Bravo, J.M. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE transactions on medical imaging*, 30(1) : 146-158 (2011).
10. Li, Q., Feng, B., Xie, L., Liang, P., Zhang, H. and Wang, T., 2016. A Cross-Modality Learning Approach for Vessel Segmentation in Retinal Images. *IEEE transactions on medical imaging*, 35(1) : 109-118 (2016).
11. Ralli, J. *PDE Based Image Diffusion and AOS*. 2014.
12. Staal, J., Abrmoff, M.D., Niemeijer, M., Viergever, M.A. and van Ginneken, B. Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging*, 23(4) : 501-509 (2004).
13. Hoover, A.D., Kouznetsova, V. and Goldbaum, M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical imaging*, 19(3), pp.203-210 (2000).
14. Soares, J.V., Leandro, J.J., Cesar, R.M., Jelinek, H.F., Cree, M.J.: Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Trans. Med. Imaging* 25(9), 1214-1222 (2006).
15. Al-Diri, Bashir, Andrew Hunter, and David Steel. An active contour model for segmenting and measuring retinal vessels. *IEEE Transactions on Medical imaging* 28(9) : 1488-1497 (2009).
16. Fraz, M.M., Barman, S.A., Remagnino, P., Hoppe, A., Basit, A., Uyyanonvara, B., Rudnicka, A.R. and Owen, C.G., An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer methods and programs in biomedicine*, 108(2) : 600-616 (2012).