Research Article Vol. 57, No. 25 / 1 September 2018 / Applied Optics 7287

applied optics

Blood vessel segmentation of fundus images via cross-modality dictionary learning

YAN YANG, FENG SHAO,* ZHENQI FU, AND RANDI FU

Faculty of Information Science and Engineering, Ningbo University, Ningbo, China
**Corresponding author: shaofeng@nbu.edu.cn*

Received 29 May 2018; revised 4 August 2018; accepted 4 August 2018; posted 6 August 2018 (Doc. ID 332882); published 28 August 2018

Automated retinal blood vessel segmentation is important for the early computer-aided diagnosis of some ophthalmological diseases and cardiovascular disorders. Traditional supervised vessel segmentation methods are usually based on pixel classification, which categorizes all pixels into vessel and non-vessel pixels. In this paper, we propose a new retinal vessel segmentation method with the motivation to extract vessels based on vessel block segmentation via cross-modality dictionary learning. For this, we first enhance the structural information of vessels using multi-scale filtering. Then, cross-modality description and segmentation dictionaries are learned to build the intrinsic relationship between the enhanced vessels and the labeled ground truth vessels for the purpose of vessel segmentation. Also, effective pre-processing and post-processing are adopted to promote the performance. Experimental results on three benchmark data sets demonstrate that the proposed method can achieve good segmentation results.

OCIS codes: (100.0100) Image processing; (170.3880) Medical and biological imaging; (100.2980) Image enhancement; (170.4470) Ophthalmology.

https://doi.org/10.1364/AO.57.007287

1. INTRODUCTION

Fundus images [1] are one type of medical images that contains retinal vessel, optic disc, macula, and other ophthalmological structures. They are widely used for the noninvasive diagnosis of ophthalmological diseases such as age-related macular degeneration [2], diabetic retinopathy [3], glaucoma [4], etc. It is known that diabetic retinopathy is the main cause for blindness [5], which is closely related to retinal vascular structures appearing as treelike branches originating from the optic disc. In addition to ophthalmological diseases, since the retinal vessel is the only vascular structure of the human blood circulation system that can be observed noninvasively [6], some cardiovascular diseases such as stroke, hypertension, and arteriosclerosis can also be diagnosed by analyzing changes of diameter, branch pattern, and tortuosity of retinal vessels. Moreover, the vessel's structural information can also be used to assist multimodal retinal image registration [7].

However, since manual annotation and measurement for vessels by human experts is an extremely time-consuming and experience-demanding task [8,9], automatic vessel segmentation that can efficiently detect vessels captured by different cameras becomes especially crucial to assist the diagnosis of various ophthalmological and cardiovascular diseases. Nevertheless, due to the distinctive characteristics of fundus images, blood vessel segmentation is not a simple image segmentation task. Although numerous attempts have been made in the area of automated fundus vessel segmentation over the last decades, this task is still active and challenging, and it can be generally categorized into two classes: unsupervised methods and supervised methods [10].

Unsupervised vessel segmentation methods do not need to train models with the pre-labeled ground truth vessels, such as filtering-based [11,12], morphological-processing-based [13], vessel-tracking-based [14] and model-based techniques [15,16]. Fraz *et al.* [17] developed a vascular tree detection method that combines vessel skeleton extraction and morphological filters to detect the vessels. Yu *et al.* [18] generated a vessel probability map using the Hessian matrix and used local second-order entropy thresholding to segment the vessels. Annunziata *et al.* [19] detected vessel structures using a multi-scale Hessian filter approach after inpainting exudates. Azzopardi *et al.* [20] introduced the combination of shifted filter responses (COSFIRE) to detect bar-shaped retinal vascular structures. Zhao *et al.* [21] combined an active contour model and compactness-based saliency detection to extract vessels after a Retinex-based vascular enhancement. However, even with their low complexity and high efficiency, the main disadvantage of the unsupervised methods is that the segmentation accuracy will be slightly decreased when applied to pathological retinal images.

On the other hand, supervised methods, which need to train a classifier to discriminate all pixels into vessel and non-vessel categories, usually produce better performance than the unsupervised methods. Using feature vectors extracted from the fundus images and their ground-truth label information, a classifier can be trained using the existing neural network (NN), support vector machine (SVM), and random forest (RF) techniques. Roychowdhury *et al.* [22] extracted two different vessels by high-pass filtering and morphological processing to extract the main vessels, and the remaining vessel pixels are classified by a Gaussian mixture model (GMM) classifier. Aslani and Sarnel [23] trained a random forests (RF) classifier with hybrid feature vectors to classify vessel and non-vessel pixels. Ricci and Perfetti [24] employed two orthogonal line detectors to construct a feature vector for the vascular classification using a SVM. Li *et al.* [25] learned a cross-modality transformation mapping function between the original retinal images and the vessel labels by deep neural network. However, the disadvantage of supervised methods is the poor generalization capability in predicting the segmented vessels across different databases.

Even though fundus blood vessel segmentation is an attractive topic, how to construct an effective and efficient segmentation solution is still an open issue. In this work, focusing on the advantages of the unsupervised and supervised methods, we try to provide a vessel segmentation framework that can achieve good segmentation accuracy and generalization capability. For this, motived by [26], we do not take segmentation as a simple pixel-wise classification problem, but attempt to establish the intrinsic relationship between the vascular enhanced images and the corresponding vessel labels by learning the description and segmentation dictionaries. In addition, due to the big difference between the thick and thin retinal vessels, extracting both of them simultaneously seems difficult, so we hereby use a filtering-based technique to enhance the vessels at different scales first and attempt to select the thin and thick vessel blocks for the cross-modality dictionary learning to further promote segmentation performance. The contributions of this work are summarized as follows:

(1) We learn description and segmentation dictionaries to establish a block-to-block relationship between the enhanced image blocks and the labeled vessel blocks. Thus, fast detection of vessel trees can be achieved via vessel block segmentation.

(2) In order to acquire more vascular details, we not only use the multi-scale filtering to enhance the vessel structure but also learn cross-modality dictionaries from the selected thin and thick vessel blocks to detect vascular structures.

(3) Comprehensive experiments are conducted on three benchmark data sets, and the results show that the proposed method can result in good segmentation performances.

In the remainder of this paper, the proposed method is described in Section 2 and the performance of our method is assessed by experiments in Section 3. Finally, conclusions are drawn in Section 4.

2. METHOD

Given an input fundus image, our method fulfills the segmentation task following three main stages: vessel enhancement, cross-modality dictionary learning, and vessel segmentation. In the vessel enhancement stage, blood vessel details are enhanced by using a multi-scale Hessian-based filtering model. In the cross-modality dictionary-learning stage, description and segmentation dictionaries are learned simultaneously from the enhanced vessel blocks (source modality) and the labeled blocks (target modality) for vessel extraction. In the vessel segmentation stage, vessels are segmented based on the learned description and segmentation dictionaries. Besides, additional pre-processing and post-processing operations are performed to ensure higher segmentation accuracy by eliminating unexpected branches. The overall architecture of the proposed method is illustrated in Fig. 1.

A. Pre-Processing

Fundus images usually suffer from inhomogeneous luminosity and varying contrast, which will largely degrade the subsequent vessel segmentation performance. In addition, the presence of

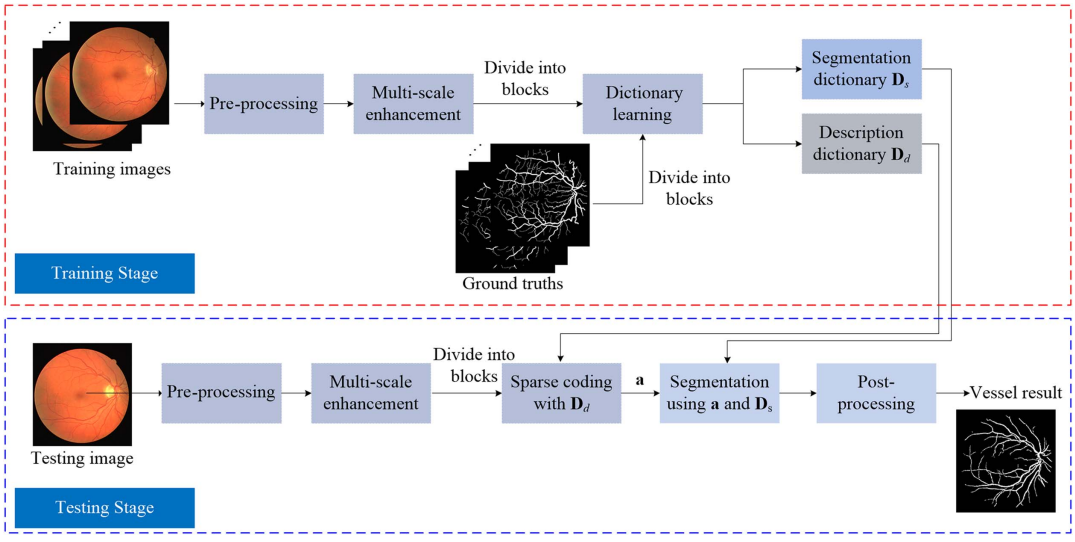


Fig. 1. Overall architecture of the proposed method.

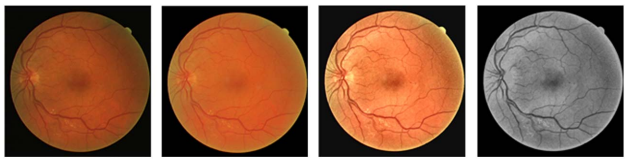


Fig. 2. Example of pre-processing (a) original image, (b) luminosity normalization, (c) contrast enhancement, (d) disturbance reduction.

an optic disc or exudates will also degrade the segmentation accuracy. To reduce erroneous detection caused by these unfavorable factors, we implement a pre-processing operation from three aspects: (1) Luminosity normalization: since the R, G, and B channels contain both luminosity information and color information, in order to enhance the luminosity and preserve the color information, we transfer the image from RGB to hue, saturation, value (HSV) color space and adjust image luminosity by gamma correction on the V channel to reduce the over-bright or dark regions and make the luminosity homogeneous. (2) Contrast enhancement: inspired by [27], we transfer the luminosity normalized images to lab color space and implement the contrast-constrained adaptive histogram equalization (CLAHE) algorithm on the L channel to enhance the contrast between the vessel and background. (3) Disturbance reduction: we apply a Gaussian filter to reduce the noises and the brightness of the optic disc or exudates via a simple thresholding. Figure 2 shows the examples of pre-processing.

B. Vessel Enhancement

Hessian-based methods have been proved effective in vessel enhancement [28], with the purpose of extracting principal directions by second-order structures. The Hessian matrix is computed as

$$H(x,y) = \begin{bmatrix} I \otimes \frac{\partial^2 G_\sigma}{\partial x^2} & I \otimes \frac{\partial^2 G_\sigma}{\partial x \partial y} \\ I \otimes \frac{\partial^2 G_\sigma}{\partial y \partial x} & I \otimes \frac{\partial^2 G_\sigma}{\partial y \partial x} \end{bmatrix}, \tag{1}$$

$$G_\sigma = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{2}$$

where σ is the standard deviation of the Gaussian convolution kernel to detect and match vessel width, and I denotes the pre-processed image.

The eigenvalues of the Hessian matrix are calculated. Here we denote the one with a bigger absolute value as λ_2 and the one with smaller value as λ_1. Since the eigenvalues of the Hessian matrix can reflect the direction of an image, a probability function is generated to detect vessels with the two eigenvalues,

$$P_v(\sigma) = \begin{cases} \exp\left(-\frac{(R_b)^2}{2\alpha^2}\right) \times \left(1 - \exp\left(-\frac{S^2}{2\beta^2}\right)\right), & \text{if } |\lambda_2| > |\lambda_1| \\ 0, & \text{otherwise} \end{cases}, \tag{3}$$

where $R_b = \lambda_1/\lambda_2$, $S = \sqrt{\lambda_1^2 + \lambda_2^2}$, α and β are the adjustable parameters, and here we determine $\alpha = 0.5$ and $\beta = 15$. P_v represents a probability function to identify whether a pixel belongs to a vessel pixel with a value varied from 0 to 1 (a large P_v means a large probability of being a vessel pixel).

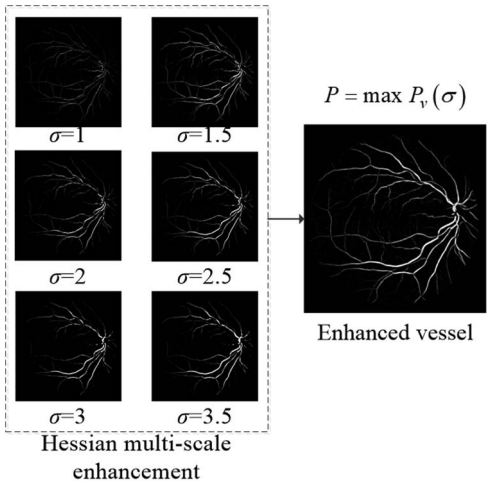


Fig. 3. Example of multi-scale vessel enhancement.

To enhance the vessels regardless their width, we calculate the probability maps under different scales σ and detect the maximum response to match the size of the vessel as

$$P = \max_{\sigma \in [\sigma_{\min}, \sigma_{\max}]} P_v(\sigma), \tag{4}$$

where $\sigma_{\min}$ and $\sigma_{\max}$ represents the minimum and maximum value of σ, and P represents the enhanced vessel map. In the experiment, we set six spatial scales with $\sigma_{\min} = 1$ and $\sigma_{\max} = 3.5$. As shown in Fig. 3, by multi-scale processing, the thick and thin vessels are simultaneously enhanced.

C. Cross-Modality Dictionary Learning

Unlike the traditional pixel-based segmentation method that usually trains a regression model to predict the pixel classification, we do not take the segmentation as a simple classification task but try to establish the intrinsic relationship via cross-modality data transformation. For this, we define the enhanced vessel images as a source modality and the ground-truth-labeled vessels as a target modality. Then, the projection relationship between the source modality and the target modality is established via cross-modality dictionary learning. Thus, information from the source modality can be easily projected to the target modality to get the classification. As shown in Fig. 4,

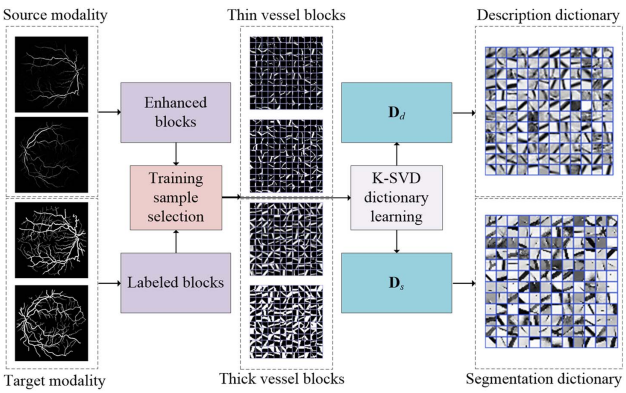


Fig. 4. Cross-modality dictionary learning.

cross-modality dictionary learning mainly contains two aspects: training sample selection and dictionary learning via K-means singular value decomposition (K-SVD) [29–31].

1. Training Sample Selection

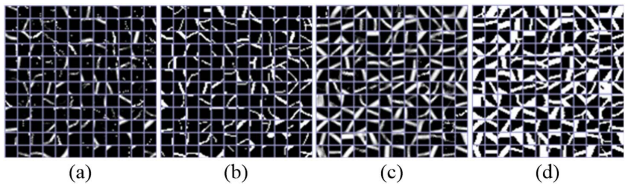
To construct the training samples, numerous overlapping blocks with a size of 10×10 are randomly selected from each training image. To select the blocks with richer vascular structure, we select blocks with large variance among all the blocks on the ground-truth-vessel image (to remove blocks without a vessel structure). The corresponding blocks from the vessel-enhanced images are also selected. In addition, to eliminate the influence of vessel thickness and to preserve the segmentation capability of the learned dictionaries, we select both thin and thick vessel blocks to construct the training samples. As a result, we obtain two types of feature vector matrices $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_P]$ and $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_P]$, where $\mathbf{x}_p \in \mathbb{R}^{n\times 1}$ and $\mathbf{y}_p \in \mathbb{R}^{n\times 1}$ represent the matched blocks in the vessel-enhanced images and ground-truth-vessel images containing n pixels and $p = 1, \ldots, P$. Here, $n = 100$ and $P = 100000$. Although the discriminative ability of our dictionaries will be increased by selecting more blocks, the training complexity will be increased correspondingly. As a trade-off, we select total of 100,000 blocks from all training images. Figure 5 shows the partial blocks selected in the training set.

2. Dictionary Learning

In this paper, the description dictionary and segmentation dictionary are learned simultaneously using the selected enhanced blocks and their corresponding labeled blocks (containing both thin and thick vessel blocks). Given a fixed least non-zero element T_0, the dictionary learning demands the smallest reconstruction error, and the objective function is formulated as follows:

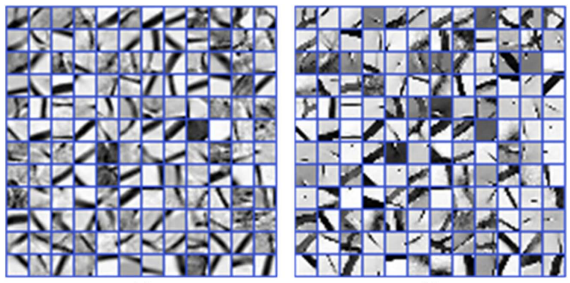
$$\langle \mathbf{D}_d, \mathbf{D}_s, \mathbf{A} \rangle = \underset{\mathbf{D}_d, \mathbf{D}_s, \mathbf{A}}{\arg\min} \|\mathbf{Y} - \mathbf{D}_d\mathbf{A}\|_2^2 + \lambda\|\mathbf{X} - \mathbf{D}_s\mathbf{A}\|_2^2 \quad \text{s.t.} \quad \forall i, \quad \|\mathbf{a}_i\|_0 \le T_0, \tag{5}$$

where $\mathbf{D}_d$ is the dictionary learned in the source modality (defined as the description dictionary), $\mathbf{D}_s$ is the dictionary learned in the target modality (defined as the segmentation dictionary), λ is a parameter to control the trade-off between the reconstruction errors in the source and target modalities, $\mathbf{A}$ is the sparse coefficient matrix, $\|\cdot\|_0$ represents the l_0 norm, which calculates the amount of non-zero values in the matrix, and $\|\cdot\|_2$ represents the l_2 norm.



(a) (b) (c) (d)

Fig. 5. Partial training samples (a) thin blocks from the enhanced images, (b) thin blocks from ground truth vessel images, (c) thick blocks from the enhanced images, (d) thick blocks from ground truth vessel images.



(a) (b)

Fig. 6. Partial atoms of (a) description dictionary and (b) segmentation dictionary.

To optimize Eq. (5), it is rewritten as

$$\langle \mathbf{D}_d, \mathbf{D}_s, \mathbf{A} \rangle = \underset{\mathbf{D}_d, \mathbf{D}_s, \mathbf{A}}{\arg\min} \left\| \begin{pmatrix} \mathbf{Y} \\ \sqrt{\lambda}\mathbf{X} \end{pmatrix} - \begin{pmatrix} \mathbf{D}_d \\ \sqrt{\lambda}\mathbf{D}_s \end{pmatrix} \mathbf{A} \right\|_2^2 \quad \text{s.t.} \quad \forall i, \quad \|\mathbf{a}_i\|_0 \le T_0. \tag{6}$$

Let $\mathbf{Y}_{\text{new}} = ((\mathbf{Y})^T, \sqrt{\lambda}(\mathbf{X})^T)^T$, $\mathbf{D}_{\text{new}} = ((\mathbf{D}_d)^T, \sqrt{\lambda}(\mathbf{D}_s)^T)^T$, and the optimization of Eq. (6) is equivalent to solving the following problem:

$$\langle \mathbf{D}_{\text{new}}, \mathbf{A} \rangle = \underset{\mathbf{D}_{\text{new}}, \mathbf{X}}{\arg\min} \|\mathbf{Y}_{\text{new}} - \mathbf{D}_{\text{new}}\mathbf{A}\|_2^2 \quad \text{s.t.} \quad \forall i, \quad \|\mathbf{a}_i\|_0 \le T_0. \tag{7}$$

Equation (7) is a standard sparse coding problem, which can be efficiently solved by the K-SVD algorithm [29–31]. Thus, the desired $\mathbf{D}_d$ and $\mathbf{D}_s$ can be separated from the trained $\mathbf{D}_{\text{new}}$. Figure 6 shows the visualized partial atoms in the description and segmentation dictionaries learned from the training data set. The number of the atoms is set to 1024. It is clear that the structures and distributions of the dictionary atoms for the two dictionaries are prominent in representing vascular details.

D. Blood Vessel Segmentation

At the testing stage, given an enhanced fundus image, after extracting non-overlapping blocks, for a block vector $\mathbf{y}_i$, we compute the sparse coefficient $\hat{\mathbf{a}}_i$ w.r.t. the learned description dictionary $\mathbf{D}_d$ by solving the following optimization function:

$$\underset{\hat{\mathbf{a}}_i}{\arg\min} \|\mathbf{y}_i - \mathbf{D}_d\hat{\mathbf{a}}_i\|_2^2, \quad \text{s.t.} \|\hat{\mathbf{a}}_i\|_0 \le T_0. \tag{8}$$

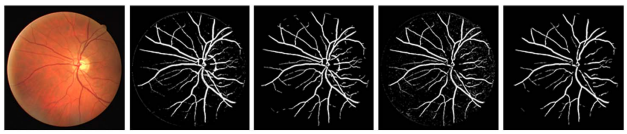
The above problems can be solved using the orthogonal matching pursuit (OMP) algorithm [32]. Then, the vessel block vector is predicted based on the learned sparse coefficient and segmentation dictionary $\mathbf{D}_s$, computed as follows:

$$\hat{\mathbf{x}}_i = \mathbf{D}_s\hat{\mathbf{a}}_i. \tag{9}$$

All pixels in the segmented vessel map are classified as vessel and non-vessel results by

$$\forall(x, y)\, V(x, y) = \begin{cases} 1, & \hat{\mathbf{x}}(x, y) \ge T_1 \\ 0, & \text{otherwise} \end{cases}, \tag{10}$$

where $V(x, y)$ is the segmented vessel, $\hat{\mathbf{x}}(x, y)$ is vessel map created by stitching all blocks, and T_1 is the threshold to obtain the binary vessel. In this experiment, we determine the optimal threshold T_1 by maximizing the accuracy rate as described in [25].



(a) (b) (c) (d) (e)

Fig. 7. Influences of pre-processing and post-processing (a) original fundus image, (b) segmentation without pre-processing and post-processing, (c) segmentation without pre-processing but with post-processing, (d) segmentation without post-processing but with pre-processing, (e) segmentation with pre-processing and post-processing.

E. Post-Processing

The obtained vessels may contain noises, small misclassified non-vessels, and undetected vessel holes. To improve segmentation accuracy, additional post-processing is implemented by the following two sub-steps: (1) removing vessel areas below 20 pixels to reduce small misclassified non-vessels and noises; (2) filling the vessel holes below 20 pixels to connect the breakpoints and holes between vessel pixels. The influences of pre-processing and post-processing on vessel segmentation are shown in Fig. 7. Obviously, pre-processing will effectively eliminate pseudo-contours, while post-processing will remove noises and fill small holes to ensure a better segmentation.

3. EXPERIMENTS

A. Training and Testing Data Sets

The proposed method relies on the training data set to learn the description dictionary and segmentation dictionary, and it uses the testing data set to further validate the performance. We select the following three databases for the experiments, as shown in Table 1.

DRIVE database [33]: contains 40 color retinal images (seven of them with pathologies) with a resolution of 584 × 565 and a 45° field of view (FOV), in which 20 images are used as the training set, and the other 20 images are used as the testing set. Two different manual annotations are provided for the testing set, and only one annotation is available in the training set. The first observer's segmentation is used as the ground truth.

STARE database [11]: contains 20 color retinal images (ten of them with pathologies) with a resolution of 700 × 605, which are captured at 35° FOV. Two different manual annotations are also provided in the database. The first observer's segmentation is used as the ground truth.

HRF database [34]: contains 45 color retinal images in which 15 images are selected from healthy patients, 15 images are selected from patients with diabetic retinopathy, and 15 images are selected from patients with glaucoma, which are captured at 60° FOV. The resolution of the images is 3504 × 2336.

Table 1. Descriptions of Three Databases

Database	Resolutions	FOV	Number	Annotation
DRIVE	584 × 565	45°	40	2
STARE	700 × 605	35°	20	2
HRF	3504 × 2336	60°	45	1

Only one ground truth segmentation is available in the database. To reduce the computational cost, all images are downsampled to a resolution of 876 × 584.

For the DRIVE database, we selected 10 enhanced images and label images from the training data set to learn the dictionaries. Since the HRF and STARE databases do not have strict training and testing sets, we hereby use the leave-one-out strategy for validation [25], in which each image in the database is tested using the dictionaries learned from the remaining images of the database.

B. Evaluation Methodology

To objectively evaluate the segmentation method, eight metrics are selected for validation: accuracy (ACC), specificity (SP), sensitivity (SE), positive predictive value (PPV), negative predictive value (NPV), F1 score (F1), G-mean (G), and Matthews correlation coefficient (MCC) [35,36]. The metrics are defined in Table 2.

$N = \mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}$, $S = (\mathrm{TP} + \mathrm{FN})/N$, and $P = (\mathrm{TP} + \mathrm{FP})/N$. True Positives (TP) is the number of pixels correctly detected as blood vessel pixels, False Negatives (FN) is the number of pixels incorrectly classified as background pixels, True Negatives (TN) is the number of pixels correctly classified as background pixels, and False Positives (FP) is the number of pixels incorrectly classified as blood vessel pixels.

Among these metrics, sensitivity (SE) measures the ratio of the correctly detected blood vessel pixels to all vessel pixels in the label image, specificity (SP) indicates the ratio of the correctly distinguished non-vessel pixels to all non-vessel pixels in the label image, and accuracy (ACC) is the ratio of the wrong classified pixels to all pixels. Positive predictive value (PPV) represents the correctly segmented vessel pixels to all the vessel pixels in the segmented image, while negative predictive value (NPV) indicates the correctly distinguished non-vessel pixels to all non-vessel pixels in the segmentation image. MCC is a correlation coefficient between the segmented vessel and labeled vessel. The F1 score is the mean value of PPV and SE with a maximum value of 1 and the lowest value 0. Similarly, the G-mean is a metric that measures the trade-off between SE and SP by taking their geometric mean.

C. Results

1. Segmentation Results

In order to evaluate the performance of our method, we carry out experiments on three databases: DRIVE, STARE, and

Table 2. Descriptions of Metrics and Calculation Formulas

Metrics	Formulas
SE	$\mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$
SP	$\mathrm{TN}/(\mathrm{TN} + \mathrm{FP})$
ACC	$(\mathrm{TN} + \mathrm{FP})/(\mathrm{TN} + \mathrm{TP} + \mathrm{FN} + \mathrm{FP})$
PPV	$\mathrm{TP}/(\mathrm{TP} + \mathrm{FP})$
NPV	$\mathrm{TN}/(\mathrm{TN} + \mathrm{FN})$
G	$\sqrt{\mathrm{SE} \times \mathrm{SP}}$
F1	$(2 \times \mathrm{PPV} \times \mathrm{SE})/(\mathrm{PPV} + \mathrm{SE})$
MCC	$(\mathrm{TP}/N - S \times P)/\sqrt{P \times S \times (1 - S)(1 - P)}$

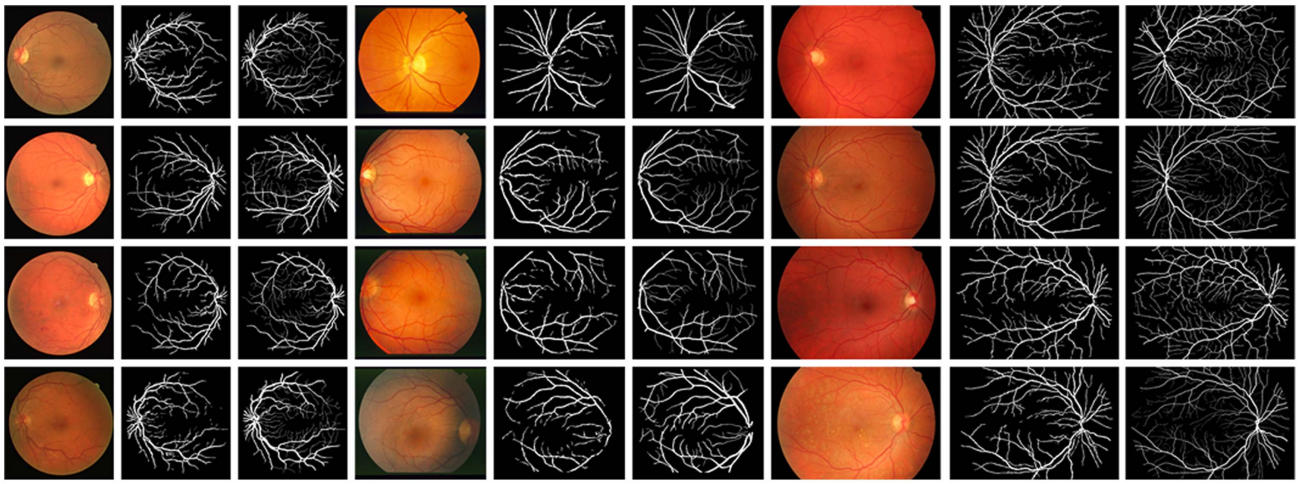


Fig. 8. Examples of vessel segmentation on the DRIVE, STARE, and HRF databases; (a), (b), and (c) are the original images, segmentation results, and manual annotations on DRIVE, respectively; (d), (e), and (f) are the original images, vessel segmentation results, and manual annotations on STARE, respectively; (g), (h), and (i) are the original images, vessel segmentation results, and manual annotations on HRF.

HRF. By comparing our segmentation results with the manual annotated vessel labels, we calculated the eight metrics introduced above based on pixel similarity. Figure 8 shows the partial segmentation results on DRIVE, STARE and HRF, and the last column in each database contains the pathological fundus images. By comparing the segmentation results obtained with the manually annotated vessel label, we can observe that our method can segment the main vessels very well and can resist the influence of pathologies, but it still has a limitation in detecting tiny vessels, especially on the DRIVE and HRF databases. Table 3 shows the segmentation results. The average ACC, SP, and SE values reach 0.9583, 0.9792, and 0.7393, respectively in the DRIVE database; 0.9549, 0.9740, and 0.7265 in the HRF database; and 0.9531, 0.9731, and 0.7046 in the STARE database, which indicates the good performance of our method.

2. Comparison with Other Methods

Tables 4–6 illustrate the segmentation performance compared with other methods on three databases, respectively. From the results, we can observe that our method achieves the highest accuracy, NPV, F1, and G values, relatively high SP, and acceptable SE, PPV, and MCC values. The comparison results show that our method has its strengths in less noises (higher SP values) and limitations in detecting tiny vessels (relatively lower SE values). We can also observe from these tables that the SE of our method is relatively low due to the limitation of training samples in reflecting thin vessels, indicating that our method still needs to extract more vessel details.

3. Cross-Database Validation

To analyze the generalization capability of our method, we further conduct cross-database validation (e.g., dictionaries trained

Table 3. Segmentation Results on DRIVE, HRF, and STARE Databases

Database	ACC	SP	SE	PPV	NPV	F1	G	MCC
DRIVE	0.9583	0.9792	0.7393	0.7770	0.9753	0.7545	0.8501	0.7340
HRF	0.9549	0.9740	0.7265	0.7003	0.9771	0.7093	0.8403	0.6873
STARE	0.9531	0.9731	0.7046	0.6984	0.9761	0.6904	0.8248	0.6716

Table 4. Performance Comparison of Different Segmentation Methods on DRIVE Database

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Fraz [37]	0.9430	0.9768	0.7152	**0.8205**	0.9587	—	—	0.7333
Marin [38]	0.9452	**0.9801**	0.7067	—	—	—	—	—
Fraz [17]	0.9422	0.9742	0.7302	0.8112	0.9600	—	—	0.7359
Vega [39]	0.9412	0.9600	**0.7444**	—	—	0.6884	—	0.6617
Palomera [40]	0.9220	0.9610	0.6600	—	—	—	—	—
Biswal [41]	0.9500	0.9700	0.7100	—	—	0.7500	0.8500	**0.7600**
Proposed	**0.9583**	0.9792	0.7393	0.7770	**0.9753**	**0.7545**	**0.8501**	0.7340

Table 5. Performance Comparison of Different Segmentation Methods on HRF Database

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Odstrcilik [34]	0.9494	0.9669	0.7741	—	—	—	—	—
Zhao [42]	0.9410	0.9420	0.7490	—	—	—	—	—
Yu [18]	0.9514	0.9685	**0.7810**	—	—	—	—	—
Proposed	**0.9549**	**0.9740**	0.7265	0.7003	0.9771	0.7093	0.8403	0.6873

Table 6. Performance Comparison of Different Segmentation Methods on STARE Database

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Fraz [37]	0.9437	0.9665	0.7409	**0.7363**	0.9709	—	—	**0.7003**
Marin [38]	0.9526	**0.9819**	0.6944	—	—	—	—	—
Fraz [17]	0.9423	0.9660	0.7318	0.7294	0.9700	—	—	0.6908
Vega [39]	0.9483	0.9671	**0.7019**	—	—	0.6616	—	0.6400
Palomera [40]	0.9240	0.9400	0.7790	—	—	—	—	—
Proposed	**0.9531**	0.9731	0.7046	0.6984	**0.9761**	**0.6904**	0.8248	0.6716

on the DRIVE database to test the HRF and STARE databases, dictionaries trained on the HRF database to test the DRIVE and STARE databases, or dictionaries trained on the STARE database to test the DRIVE and HRF databases). The cross-database validation results are illustrated in Table 7. Observed from the table, the selection of the training samples will have a large influence on the segmentation results. For example, the ACC, SP, and SE metrics on the STARE database are largely decreased when trained on the HRF database. The reason may be that the HRF database has thinner vessel types. We can also observe from the table that the dictionaries trained from the STARE database are not suitable for vessel segmentation on the DRIVE and HRF databases. The reason may be that the STARE database contains many pathological or low-quality images, and the ground truth vessel thickness is quite different from those of the HRF and DRIVE databases. In other words, by carefully selecting the training samples that include the vessel types in the testing database, the segmentation performance of the proposed method is usually stable.

D. Further Discussion

In this paper, we attempt to learn cross-modality dictionaries for retinal vessel segmentation. Although our model demonstrates its strength in vessel detection for different types of retinal images, the following issues still deserve to be considered:

(1) The proposed vessel segmentation method is highly dependent on the samples selected for dictionary learning. In this paper, we only select samples from thin and thick vessel blocks, but the segmentation performance on thin vessels is not as satisfactory as expected. In future work, we consider adding samples with different thicknesses. In addition, more types of images should be considered, such as OCT images [43,44].

(2) As illustrated in Fig. 7, the pre-processing and post-processing can decrease small vessel holes and noises, especially in the vicinity of the optic disc, but the vascular details may be also removed, which indicates a demand for better pre-processing and post-processing approaches.

(3) For our block-based segmentation model, in addition to the Hessian-based multi-scale enhancement feature, we expect to promote the efficiency and robustness by adopting richer enhanced vessel features and better supervised dictionary learning models, such as label-consistent k-means singular value decomposition (LC-KSVD) [45–47], which can add labels for vessels of different thickness.

(4) In the vessel enhancement stage, a large scale factor for the big vessel may make the small vessel fat and affect its morphological information. This may lead to fat segmentation results, especially for the thin vessels. Therefore, a better enhancement method will be considered in the future.

(5) Similar to the existing supervised methods, our method also needs manual labeled vessels for training. Thus, if one database does not have enough labels for training, the availability of our method on the database will be limited. Therefore, cross-database capability becomes very important for the supervised methods.

(6) Compared with neural-network-based supervised methods, our dictionary-learning-based method can only segment block-wise vessels, but the training process is relatively simple and does not need a complex parameter selection process.

Table 7. Performance of Cross-Database Validation

Database	ACC	SP	SE	PPV	NPV	F1	G	MCC
DRIVE (trained on HRF)	0.9504	0.9723	0.7232	0.7210	0.9734	0.7167	0.8373	0.6928
DRIVE (trained on STARE)	0.9462	0.9839	0.5547	0.6250	0.7697	0.6421	0.7378	0.6250
STARE (trained on DRIVE)	0.9521	0.9772	0.6348	0.7256	0.9713	0.6518	0.7782	0.6413
STARE (trained on HRF)	0.9494	0.9685	0.7040	0.6789	0.9768	0.6657	0.8171	0.6528
HRF (trained on DRIVE)	0.9445	0.9644	0.7022	0.6307	0.9752	0.6581	0.8215	0.6330
HRF (trained on STARE)	0.9400	0.9700	0.5850	0.6390	0.9654	0.5979	0.7513	0.5740

4. CONCLUSIONS

In this paper, we present a new fundus vessel segmentation method that uses multi-scaling filtering to enhance the structure information of vessels and segments of retinal vessels based on vessel block segmentation instead of pixel-wise classification via cross-modality dictionary learning. As a result, our method yields a segmentation result with the advantages of good segmentation accuracy and generalization capability. As for our future work, based on this work, we plan to add richer enhanced features and add labels for vessels of different thicknesses to promote segmentation accuracy.

Funding. National Natural Science Foundation of China (NSFC) (61622109); Natural Science Foundation of Zhejiang Province of China (R18F010008); Natural Science Foundation of Ningbo (2017A610112). It was also sponsored by the K. C. Wong Magna Fund at Ningbo University.

REFERENCES

1. M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," IEEE Rev. Biomed. Eng. **3**, 169–208 (2010).
2. R. D. Jager, W. F. Mieler, and J. W. Miller, "Age-related macular degeneration," N. Engl. J. Med. **358**, 2606–2617 (2008).
3. N. Cheung, P. Mitchell, and T. Y. Wong, "Diabetic retinopathy," Lancet **376**, 124–136 (2010).
4. J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N. M. Tan, and T. Y. Wong, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," IEEE Trans. Med. Imaging **32**, 1019–1032 (2013).
5. M. D. Abràmoff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell, and M. Lamard, "Automated analysis of retinal images for detection of referable diabetic retinopathy," JAMA Ophthalmol. **131**, 351–357 (2013).
6. S. C. Cheng and Y. M. Huang, "A novel approach to diagnose diabetes based on the fractal characteristics of retinal images," IEEE Trans. Inf. Technol. Biomed. **7**, 163–170 (2003).
7. F. Zana and J. C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," IEEE Trans. Med. Imaging **18**, 419–428 (1999).
8. T. Y. Wong, F. A. Islam, R. Klein, B. E. Klein, M. F. Cotch, C. Castro, and E. Shahar, "Retinal vascular caliber, cardiovascular risk factors, and inflammation: the multi-ethnic study of atherosclerosis (MESA)," Invest. Ophthalmol. Visual Sci. **47**, 2341–2350 (2006).
9. H. Li, W. Hsu, M. L. Lee, and T. Y. Wong, "Automatic grading of retinal vessel caliber," IEEE Trans. Biomed. Eng. **52**, 1352–1355 (2005).
10. M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, "Blood vessel segmentation methodologies in retinal images: a survey," Comput. Methods Programs Biomed. **108**, 407–433 (2012).
11. A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Trans. Med. Imaging **19**, 203–210 (2000).
12. B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," Comput. Biol. Med. **40**, 438–445 (2010).
13. A. Mendonca and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," IEEE Trans. Med. Imaging **25**, 1200–1213 (2006).
14. I. Liu and Y. Sun, "Recursive tracking of vascular networks in angiograms based on the detection deletion scheme," IEEE Trans. Med. Imaging **12**, 334–341 (1993).
15. B. Al-Diri, A. Hunter, and D. Steel, "An active contour model for segmenting and measuring retinal vessels," IEEE Trans. Med. Imaging **28**, 1488–1497 (2009).
16. S. Cetin, A. Demir, A. Yezzi, M. Degertekin, and G. Unal, "Vessel tractography using an intensity based tensor model with branch detection," IEEE Trans. Med. Imaging **32**, 348–363 (2013).
17. M. M. Fraz, A. Basit, and S. A. Barman, "Application of morphological bit planes in retinal blood vessel extraction," IET Image Proc. **26**, 373–383 (2013).
18. H. Yu, S. Barriga, C. Agurto, G. Zamora, W. Bauman, and P. Soliz, "Fast vessel segmentation in retinal images using multiscale enhancement and second-order local entropy," Proc. SPIE **8315**, 83151B (2012).
19. R. Annunziata, A. Garzelli, L. Ballerini, A. Mecocci, and E. Trucco, "Leveraging multiscale Hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation," IEEE J. Biomed. Health Inf. **20**, 1129–1138 (2016).
20. G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, "Trainable COSFIRE filters for vessel delineation with application to retinal images," Med. Image Anal. **19**, 46–57 (2015).
21. Y. Zhao, J. Zhao, J. Yang, Y. Liu, Y. Zhao, Y. Zheng, L. Xia, and Y. Wang, "Saliency driven vasculature segmentation with infinite perimeter active contour model," Neurocomputing **259**, 201–209 (2017).
22. S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Blood vessel segmentation of fundus images by major vessel extraction and subimage classification," IEEE J. Biomed. Health Inf. **19**, 1118–1128 (2015).
23. S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," Biomed. Signal Process. Control **30**, 1–12 (2016).
24. E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," IEEE Trans. Med. Imaging **26**, 1357–1365 (2007).
25. Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," IEEE Trans. Med. Imaging **35**, 109–118 (2016).
26. B. Chen, Y. Chen, Z. Shao, and L. Luo, "Blood vessel enhancement via multi-dictionary and sparse coding," Neurocomputing **200**, 110–117 (2016).
27. M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color retinal image enhancement based on luminosity and contrast adjustment," IEEE Trans. Biomed. Eng. **65**, 521–527 (2018).
28. A. Frangi, W. Niessen, K. Vincken, and M. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (1998), Vol. **1496**, pp. 130–137.
29. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. Signal Process. **54**, 4311–4322 (2006).
30. Q. Jiang, F. Shao, W. Lin, and G. Jiang, "Learning sparse representation for objective image retargeting quality assessment," IEEE Trans. Cybern. **48**, 1276–1289 (2018).
31. F. Shao, W. Tian, W. Lin, G. Jiang, and Q. Dai, "Learning sparse representation for no-reference quality assessment of multiply-distorted stereoscopic images," IEEE Trans. Multimedia **19**, 1821–1836 (2017).
32. Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of 27th Asilomar Conference on Signals, Systems and Computers* (1993), Vol. **1**, pp. 40–44.
33. J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Trans. Med. Imaging **23**, 501–509 (2004).
34. J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," IET Image Process. **7**, 373–383 (2013).
35. J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," IEEE Trans. Med. Imaging **64**, 16–27 (2017).
36. J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluim, R. Duits, and B. M. H. Romeny, "Robust retinal vessel segmentation via locally

adaptive derivative frames in orientation scores," IEEE Trans. Med. Imaging **35**, 2631–2644 (2016).
37. M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, "Retinal vessel extraction using first-order derivative of Gaussian and morphological processing," in *Advances in Visual Computing* (2011), Vol. **6938**, pp. 410–420.
38. D. Marin, A. Aquino, M. E. Gegundez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," IEEE Trans. Med. Imaging **30**, 146–158 (2011).
39. R. Vega, G. Sanchez-Ante, L. E. Falcon-Morales, S. Sossa, and E. Guevara, "Retinal vessel extraction using lattice neural networks with dendritic processing," Comput. Biol. Med. **58**, 20–30 (2015).
40. M. Palomera-Prez, M. Martinez-Perez, H. Bentez-Prez, and J. Ortega-Arjona, "Parallel multiscale feature extraction and region growing: application in retinal blood vessel detection," IEEE Trans. Inf. Technol. Biomed. **14**, 500–506 (2010).
41. B. Biswal, T. Pooja, and N. Bala Subrahmanyam, "Robust retinal blood vessel segmentation using line detectors with multiple masks," IET Image Process. **12**, 389–399 (2018).
42. Y. Zhao, Y. Zheng, Y. Liu, Y. Zhao, L. Luo, S. Yang, and J. Liu, "Automatic 2-D/3-D vessel enhancement in multiple modality images using a weighted symmetry filter," IEEE Trans. Med. Imaging **37**, 438–450 (2018).
43. A. Li, J. You, C. Du, and Y. Pan, "Automated segmentation and quantification of OCT angiography for tracking angiogenesis progression," Biomed. Opt. Express **8**, 5604–5616 (2017).
44. Z. Chu, J. Lin, C. Gao, C. Xin, Q. Zhang, C. L. Chen, L. Roisman, G. Gregori, P. J. Rosenfeld, and R. K. Wang, "Quantitative assessment of the retinal microvasculature using optical coherence tomography angiography," J. Biomed. Opt. **21**, 066008 (2016).
45. Z. Jiang, Z. Lin, and L. Davis, "Label consistent K-SVD: learning a discriminative dictionary for recognition," IEEE Trans. Pattern Anal. Mach. Intell. **35**, 2651–2664 (2013).
46. Q. Jiang, F. Shao, W. Lin, K. Gu, G. Jiang, and H. Sun, "Optimizing multi-stage discriminative dictionaries for blind image quality assessment," IEEE Trans. Multimedia **20**, 2035–2048 (2018).
47. L. Zhang, L. Li, H. Li, and M. Yang, "3D ear identification using block-wise statistics based features and LC-KSVD," IEEE Trans. Multimedia **18**, 1531–1541 (2016).