



# Discriminative dictionary learning for retinal vessel segmentation using fusion of multiple features

Yan Yang<sup>1</sup> · Feng Shao<sup>1</sup> · Zhenqi Fu<sup>1</sup> · Randi Fu<sup>1</sup>

Received: 26 November 2018 / Revised: 4 May 2019 / Accepted: 20 May 2019 / Published online: 29 May 2019  
© Springer-Verlag London Ltd., part of Springer Nature 2019

## Abstract

In recent years, automated retinal vessel segmentation has become especially essential for the early detection of some ophthalmological and cardiovascular diseases. In this paper, we have presented a new retinal vessel segmentation method via discriminative dictionary learning using fusion of multiple features, which is able to capture both thick and thin vessel structures. In the training stage, we employ six different enhancement algorithms to obtain multiple complementary features that contain rich vascular information. Then, the manually annotated ground-truth vessels are classified into thick or thin vessels as the label information, and the label consistent KSVD based framework is applied to train the dictionary for vessel segmentation. In the testing stage, comprehensive experiments are conducted on three datasets to measure segmentation performance with eight representative evaluation metrics. The average sensitivity reaches 0.7915, 0.7560 and 0.7202 respectively, suggesting that our method can segment tiny vascular structures well.

**Keywords** Retinal image · Blood vessel segmentation · Vascular enhancement · Dictionary learning

## 1 Introduction

Retinal image, also named as fundus image, is a kind of important medical images captured by fundus cameras [1]. It can provide abundant information for the diagnosis of ophthalmological diseases, such as age-related macular degeneration, glaucoma and diabetic retinopathy. Retinal vessels, appeared as treelike branches, are the only blood vessel structures in human blood vessel circulation system that can be observed noninvasively. Therefore, exudation, hemorrhage and morphology changes in retinal vessels can reflect symptoms of some systemic cardiovascular diseases, which indicates retinal vessel structures are especially important for the diagnoses of fundus diseases and some cardiovascular diseases. In addition, retinal vessels can be used in some special applications, such as multimodal image registration and optic disc localization. However, most of the retinal vessels are currently obtained by the manual measurements and annotations, which is extremely tedious and time-consuming.

Thus, it is necessary to achieve automatic blood vessel segmentation for retinal images.

Over the last decades, researchers have proposed many automatic segmentation methods, which can be broadly divided into two categories: unsupervised methods and supervised methods [2]. Unsupervised methods do not rely on ground truth information, which are relatively simple with lower complexity compared with those supervised methods. Annunziata et al. [3] segmented vessels by a multiscale Hessian filtering method after removing undesired exudates. Yu et al. [4] constructed a vessel probability map and employed a local second-order entropy thresholding method to extract the vessel structures. Azzopardi et al. [5] adopted the combination of shifted filter responses (COSFIRE) algorithm to extract the bar-shaped retinal vessel structures. Zhao et al. [6] introduced the active contour model and compactness-based saliency detection technology to detect vessels after vascular enhancement. Fraz et al. [7] extracted retinal vessels by Gaussian filtering and morphological bit plane slicing. Zhao et al. [8] adopted the weighted symmetric filters to segment vessel structures. Zhang et al. [9] used wavelet transformation to lift image dimensions and employed Gaussian filtering to enhance vessel structures.

On the other hand, supervised methods need to train models with pre-labeled ground-truth vessel annotations, which

✉ Feng Shao  
shaofeng@nbu.edu.cn

<sup>1</sup> Faculty of Information Science and Engineering, Ningbo University, Ningbo, China

can be further categorized into shallow learning approaches and deep learning approaches. The shallow learning methods mainly segment vessels by training a binary classifier to discriminate all pixels into vessel and non-vessel background pixels with the previously extracted features. The classifiers are trained with techniques like k-nearest neighbor (KNN), support vector machine (SVM), random forest (RF). Soares et al. [10] trained a Bayes classifier with the Gabor and intensity features. Fraz et al. [11] adopted the bagged decision tree to segment vessels with multiple features. Aslani et al. [12] trained the random forest classifier using different hybrid features to classify vessel pixels and non-vessel pixels. Roychowdhury et al. [13] extracted major vessel structures first and then trained the Gaussian mixture model to segment tiny vessels. Marin et al. [14] learned a neural network with gray-level and moment invariants-based features. Besides, segmentation methods based on deep learning are gradually developed. These methods do not need to manually extract features and is more intelligent than shallow learning methods. Orando et al. [15] adopted fully connected conditional random field model to detect vessels. Li et al. [16] trained a cross-modality transformation mapping function by deep learning networks for segmentation. Fu et al. [17] formulated segmentation as a boundary detection problem and used a fully convolutional neural network to solve. Yan et al. [18] segmented vessels by deep learning with joint segment-level and pixel-wise losses. Compared with unsupervised methods, supervised approaches are relatively complex but normally having a higher segmentation accuracy.

However, there are still limitations in accurate vessel extraction due to the following reasons: (1) presence of optic disk and some pathological structures like hemorrhages, lesions and exudates; (2) variability of vascular thickness and length; (3) low contrast and poor quality of some retinal images [19]; (4) central reflex on some large vessel structures; (5) inhomogeneous illumination level of retinal vessels in different local areas. In our previous work [20], a cross-modality dictionary learning framework by KSVD is established, in which description and segmentation dictionaries are learned to build intrinsic relationship between the enhanced vessels and the ground truth vessels. In this paper, we complement our previous work by further considering enriched features. We also modify the dictionary learning framework by adding a classification label term. Our method performs much better than previous work for the following contributions: (1) We adopt different enhancement algorithms to obtain complementary feature maps to comprehensively represent retinal vessel characteristics and enhance the contrast of original retinal images. (2) Considering variations in vessel thickness, we classify all vessel branches into thick or thin vessels, and assign a classification label for each block to supervise dictionary training. (3) We explore LC-KSVD based classification framework to establish the intrinsic relationship

between multiple features and the ground truths. We used LC-KSVD rather than usual KSVD mainly because it can introduce a new discriminative label constraint to supervise the training.

The rest of this paper is organized as follows. Section 2 shows the details of our method. Section 3 presents our results, and Sect. 4 gives the conclusions.

## 2 Method

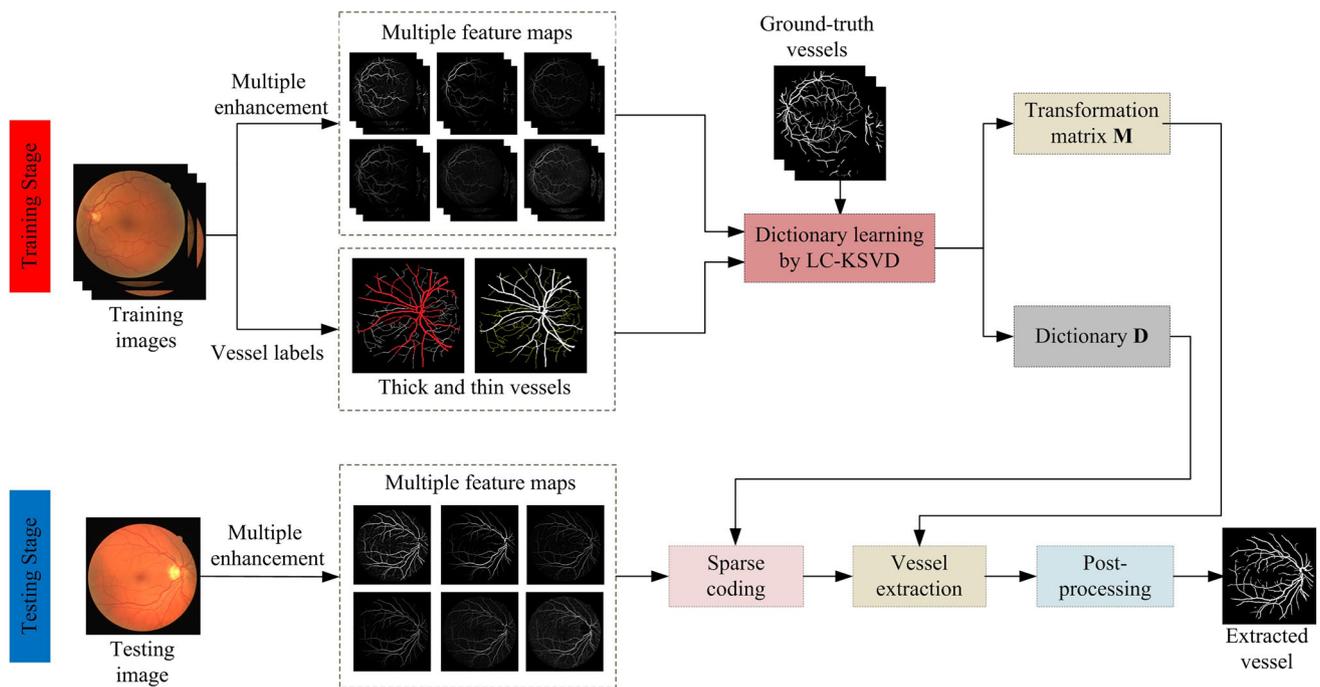
Figure 1 illustrates the overall framework of our proposed method, which consists of four main steps, feature representation, vessel classification, dictionary learning and vessel segmentation. All these steps belong to two stages: training stage and testing stage. In the training stage, LC-KSVD model is used to learn the dictionary and transformation matrix. In the testing stage, vessels are segmented with the learned dictionary and transformation matrix.

### 2.1 Feature representation

Since retinal images usually suffer from inhomogeneous luminosity, a pre-processing operation is conducted. We transfer the original image into HSV color space and perform the gamma correction on channel V to normalize the luminosity to reduce the over dark or over bright areas. Then, we transfer the image back to RGB color space and choose green channel for the consequent feature representation.

Retinal images are sometimes of low contrast and the modality of source retinal vessels is quite different from ground truths, which will degrade the performance of segmentation. Thus, instead of directly using preprocessed images for training, we adopt the enhanced feature maps with brighter vessel structures and darker background information to train the dictionary. In addition, to present and reflect vascular characteristics comprehensively, we applied six state-of-the-art enhancement algorithms, i.e., total variation by  $l_1$  norm [21], Hessian filtering [22], line structure detection, bottom hat transformation, Retinex [23] and contrast limited adaptive histogram equalization (CLAHE) [24], to acquire feature maps. These maps can well present the geometry, texture and space information of retinal vessels. The corresponding implementation details are as follows.

1. Total variation by  $l_1$  norm: TV-L1 (total variation by  $l_1$  norm) model is adopted to decompose the retinal image into smooth cartoon part and rough texture part. Here we select the texture part and transform it into a gray image as the feature map.

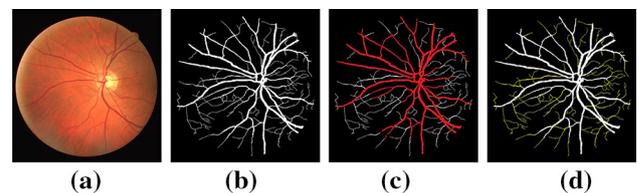


**Fig. 1** Overall framework of the proposed method

2. Hessian filtering: eigenvalues of Hessian filtered retinal image are used to generate the probability function. By adjusting the filter parameters, vessels of different thickness can be detected as the feature map.
3. Line structure detection: bottom and top hat transformations of different line structure elements are applied on original images respectively. Then the feature map is acquired by subtracting the top hat transformation map from that of bottom hat transformation.
4. Bottom hat transformation: the morphological dilation and erosion operations are applied on retinal images with a disk structural element. Then, by subtracting the original retinal images from the obtained images, we get the feature map.
5. Retinex: illumination part of retinal image is estimated first. Then, the illumination part is removed from the original image so that reflection part containing essential details can be obtained as the feature map.
6. CLAHE: we transfer retinal images into Lab color space, and implement CLAHE algorithm on channel L to enhance the contrast between vessel structures and the background.

## 2.2 Vessel classification

For LC-KSVD framework [25], classification labels for input samples should be assigned to learn a discriminative dictionary. By analyzing retinal images, we found that the majority of vessel pixels belong to thick vessels while oth-



**Fig. 2** Vessel classification **a** original image, **b** ground-truth vessels, **c** thick vessel structures, **d** thin vessels

ers belong to thin vessels. But the existing machine learning based methods treat all vessel pixels with equal importance, which may lead to poor accuracy in detecting thin vessels. Based on this, we segment the ground-truth annotations into thick and thin vessels. As shown in Fig. 2, by classifying all annotated ground-truth vessels into thick vessel structures (shown in red in Fig. 2c) or thin vessel structures (shown in yellow in Fig. 2d), a vessel thickness classification label can be assigned for each image block.

## 2.3 Dictionary learning

By using the feature representation methods, any input retinal image can be represented by a feature vector. To construct training samples for the LC-KSVD based dictionary learning, here we divide feature maps and the corresponding ground-truth images into numerous overlapping blocks with a size of  $10 \times 10$ . Then training samples, including both thin and

thick vessel blocks, are randomly selected. For the  $i$ -th sample block, we can construct a set of unique feature vector  $\mathbf{x}_i$ , ground-truth vector  $\mathbf{y}_i$  and label vector  $\mathbf{z}_i$ . Here, six feature blocks are directly stacked to construct  $\mathbf{x}_i$  for training so that the cross-modality characteristics of these feature maps can be fused for segmentation. With these vectors, we further construct the feature matrix  $\mathbf{X}$ , ground-truth matrix  $\mathbf{Y}$ , and classification matrix  $\mathbf{Z}$  for training.

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathfrak{R}^{6n \times N} \quad (1)$$

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N] \in \mathfrak{R}^{n \times N} \quad (2)$$

$$\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N] \in \mathfrak{R}^{2 \times N} \quad (3)$$

where  $n$  is the number of pixels in a block,  $n = 100$  in experiment, and  $N$  is the amount of selected sample blocks. For classification matrix  $\mathbf{Z}$ , different elements indicate different vessel thickness with a column vector  $\mathbf{z}_i = [0, 1]^T$  (thin vessel) or  $\mathbf{z}_i = [1, 0]^T$  (thick vessel).

To learn the dictionary and transformation matrix, the LC-KSVD based training model is expressed as:

$$\begin{aligned} &\langle \mathbf{D}, \mathbf{M}, \mathbf{H}, \mathbf{A} \rangle \\ &= \arg \min_{\mathbf{D}, \mathbf{M}, \mathbf{H}, \mathbf{A}} \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_2^2 + \alpha \|\mathbf{Y} - \mathbf{M}\mathbf{A}\|_2^2 + \beta \|\mathbf{Z} - \mathbf{H}\mathbf{A}\|_2^2 \\ &\quad \text{s.t. } \forall i, \|\mathbf{a}_i\| \leq T_0 \end{aligned} \quad (4)$$

where  $\mathbf{D} \in \mathfrak{R}^{6n \times K}$  is the learned dictionary,  $\mathbf{A} \in \mathfrak{R}^{K \times N}$  is the sparse matrix,  $\mathbf{H} \in \mathfrak{R}^{2 \times K}$  is the vessel thickness classification matrix, and  $\mathbf{M} \in \mathfrak{R}^{n \times K}$  is a linear transformation matrix that transforms features into vessel images. The first term  $\|\mathbf{X} - \mathbf{D}\mathbf{A}\|_2^2$  is the reconstruction error of feature maps. The second term  $\|\mathbf{Y} - \mathbf{M}\mathbf{A}\|_2^2$  is the reconstruction error of ground-truth vessels. The third term  $\|\mathbf{Z} - \mathbf{H}\mathbf{A}\|_2^2$  is the thickness classification error. In these matrices,  $K$  is the number of dictionary atoms,  $\alpha$  and  $\beta$  are parameters to control the contribution of these terms,  $\mathbf{a}_i$  represents the  $i$ -th sparse coefficient in  $\mathbf{A}$ , and  $T_0$  is a fixed least nonzero element to stop the iteration.

To optimize Eq. (4), we construct  $\mathbf{X}_{\text{new}}$  and  $\mathbf{D}_{\text{new}}$ :

$$\mathbf{X}_{\text{new}} = (\mathbf{X}^T, \sqrt{\alpha}\mathbf{Y}^T, \sqrt{\beta}\mathbf{Z}^T)^T \quad (5)$$

$$\mathbf{D}_{\text{new}} = (\mathbf{D}^T, \sqrt{\alpha}\mathbf{M}^T, \sqrt{\beta}\mathbf{H}^T)^T \quad (6)$$

Then, Eq. (4) is equivalent to solve the following formula:

$$\begin{aligned} \langle \mathbf{D}_{\text{new}}, \mathbf{A} \rangle &= \arg \min_{\mathbf{D}_{\text{new}}, \mathbf{A}} \|\mathbf{X}_{\text{new}} - \mathbf{D}_{\text{new}}\|_2^2 \\ &\quad \text{s.t. } \forall i, \|\mathbf{a}_i\| \leq T_0 \end{aligned} \quad (7)$$

Equation (7) is a standard sparse coding problem, which can be efficiently solved by KSVD algorithm. Thus, the desired  $\mathbf{D}$  and  $\mathbf{M}$  can be easily separated from  $\mathbf{D}_{\text{new}}$ .

## 2.4 Vessel segmentation

At the vessel segmentation stage, given an original retinal image, we use the same pre-processing and feature extraction methods to construct feature vectors for non-overlapping image blocks. And sparse coefficients for reconstructing a feature vector  $\hat{\mathbf{x}}_i$  can be obtained by solving the following optimization function, w.r.t. the learned dictionary  $\mathbf{D}$ :

$$\arg \min_{\hat{\mathbf{a}}_i} \|\hat{\mathbf{x}}_i - \mathbf{D}\hat{\mathbf{a}}_i\|_2^2, \quad \text{s.t. } \forall i, \|\hat{\mathbf{a}}_i\| \leq T_0 \quad (8)$$

Equation (8) is efficiently solved using the batch orthogonal matching pursuit (OMP) algorithm. Then, the corresponding vessel vector  $\mathbf{v}_i$  can be acquired with the transformation matrix  $\mathbf{M}$  and the obtained sparse coefficients, which can be calculated as:

$$\mathbf{v}_i = \mathbf{M}\hat{\mathbf{a}}_i \quad (9)$$

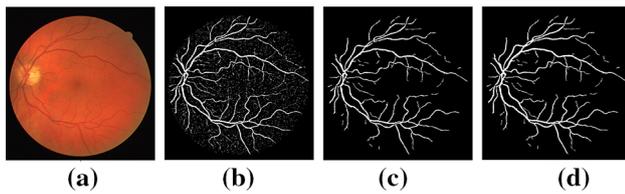
Then, all pixels in the segmented vessel map are classified into vessel and non-vessel pixels by:

$$\mathbf{V}(x, y) = \begin{cases} 1, & \mathbf{v}(x, y) \geq T_1 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

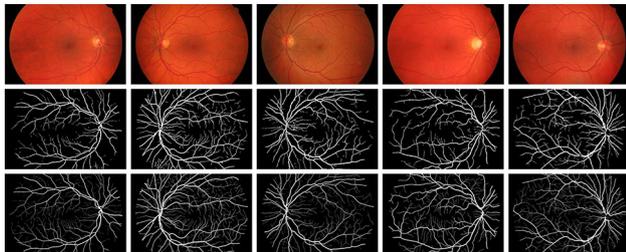
where  $\mathbf{V}(x, y)$  is the segmented vessel, and  $T_1$  is the threshold to obtain the binary vessel. In this experiment, we determine the optimal threshold  $T_1$  by maximizing the accuracy value as described in paper [15].

## 2.5 Post-processing

Since the obtained vessels may contain noises, small misclassified non-vessel structures and some undetected vessel holes, to reduce errors caused by these factors and get smooth segmentation results, a post-processing operation is performed with two steps: (1) Noise removal: connected areas below 20 pixels (determined by experiments) are removed to decrease noises and tiny misclassified non-vessel structures. (2) Hole-filling: connected holes below 20 pixels are filled to connect breakpoints and tiny holes between vascular pixels. An example of post-processing operation is shown in Fig. 3.



**Fig. 3** Post-processing, **a** original image, **b** initially extracted vessels, **c** noise removal, **d** hole-filling



**Fig. 4** Segmentation results on HRF, the first row shows original images, the second row shows segmentation images, and the third row shows manual annotations

## 3 Experiments

### 3.1 Databases

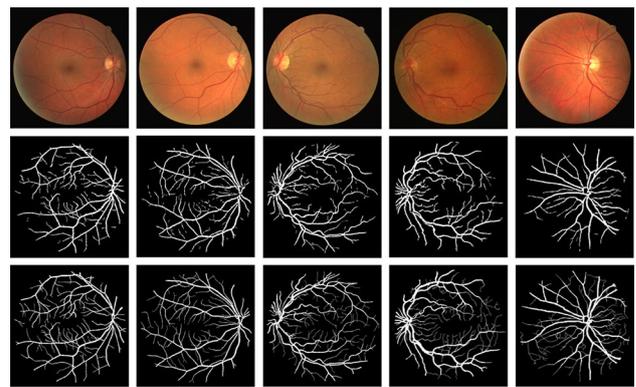
We conduct experiments on HRF [26], DRIVE [27] and STARE [28] for performance evaluation. All evaluation metrics are calculated within the field of view (FOV).

HRF dataset contains 45 color retinal images with a resolution of  $3504 \times 2336$  and at  $60^\circ$  FOV, in which 15 images are from healthy patients, 15 images are from patients with diabetic retinopathy, and the rest 15 images are from patients with glaucoma. Only one ground truth is provided by a group of experts. Since the resolution of the image is relatively high, to reduce the computational cost of our method, all images are down-sampled to a resolution of  $876 \times 584$ .

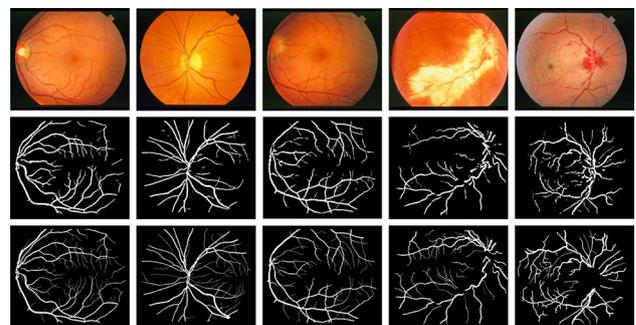
DRIVE dataset is composed of 40 retinal images captured at  $45^\circ$  FOV with a resolution of  $584 \times 565$ . The database is composed of a training and testing set, both containing 20 images. Two different benchmark annotations are provided in testing set, but only one manual annotation is available in training set. Thus, the first annotation is selected as the ground truth.

STARE dataset contains 20 retinal images (ten of them containing pathologies) captured at  $35^\circ$  FOV with a resolution of  $700 \times 605$ . Two different manual annotations are provided for each image. In the experiment, the first annotation is chosen as the ground truth.

Since HRF and STARE have no strict training and testing set, we employ leave-one-out cross validation. For DRIVE, we select 15 representative images from the training set to learn the dictionary and transformation matrix, and test all images in the testing set.



**Fig. 5** Segmentation results on DRIVE, the first row shows original images, the second row shows segmentation images, and the third row shows the manual annotations



**Fig. 6** Segmentation results on STARE, the first row shows original images, the second row shows segmentation images and the third shows manual annotations

### 3.2 Evaluation methodology

The performance is analyzed by comparing the segmentation results with ground truths. Eight different evaluation metrics are used [15], all of them are computed from true positives TP, true negatives TN, false positives FP, and false negatives FN. Among these metrics, accuracy (ACC) measures the ratio of correctly identified pixels against all pixels, specificity (SP) is the ratio of pixels correctly classified as non-vessel against all non-vessel pixels in the ground truth, while sensitivity (SE) represents the ratio of pixels that are correctly classified as vessel against all vessel pixels in ground truth. Positive predictive value (PPV) is the ratio of pixels classified as vessel that are correctly identified, while negative predictive value (NPV) represents the ratio of pixels classified as non-vessel that are correctly identified. F1-score is the weighted mean value of PPV and SE to measure the trade-off between PPV and SE. Similarly, G-mean (G) is to evaluate the trade-off between SP and SE. Matthews correlation coefficient (MCC) is the correlation coefficient between segmented vessels and the ground truths. Let  $S = (TP+FN)/M$ ,  $P = (TP+FP)/M$  and  $M = TP+TN+FP+FN$ . These evaluation metrics can be

**Table 1** Performance comparison on HRF

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Odstrcilik [26]	0.9494	0.9669	0.7741	–	–	–	0.8651	–
Zhao [8]	0.9410	0.9420	0.7490	–	–	–	0.8400	–
Yu [4]	0.9514	0.9685	0.7810	–	–	–	0.8697	–
Orlando [15]	–	0.9584	0.7874	0.6630	–	0.7158	0.8686	–
Yan [18]	0.9437	0.9592	0.7881	0.6647	–	–	0.8695	–
Work [20]	0.9467	0.9688	0.7265	0.7003	0.9725	0.7094	0.8381	0.6825
Proposed	0.9517	0.9676	0.7915	0.7079	0.9790	0.7449	0.8745	0.7125

**Table 2** Performance comparison on DRIVE (2nd means the second manual annotation of DRIVE)

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Fraz [7]	0.9422	0.9742	0.7302	0.8112	0.9600	–	0.8434	0.7359
Fraz [29]	0.9430	0.9768	0.7152	0.8205	0.9587	–	0.8358	0.7333
Vega [30]	0.9412	0.9600	0.7444	–	–	0.6884	0.8454	0.6617
You [31]	0.9434	0.9751	0.7410	–	–	–	0.8500	–
Al-Diri [32]	0.9452	0.9551	0.7282	–	–	–	0.8340	–
Orlando [15]	–	0.9684	0.7897	0.7854	–	0.7857	0.8741	0.7556
2nd	0.9477	0.9744	0.7665	0.8162	0.9661	0.7873	0.8634	0.7601
Work [20]	0.9401	0.9689	0.7373	0.7770	0.9628	0.7545	0.8483	0.7258
Proposed	0.9421	0.9696	0.7560	0.7854	0.9644	0.7673	0.8555	0.7365

**Table 3** Performance comparison on STARE (2nd means the second manual annotation of STARE)

Method	ACC	SP	SE	PPV	NPV	F1	G	MCC
Fraz [7]	0.9423	0.9660	0.7318	0.7294	0.9700	–	0.8408	0.6908
Fraz [29]	0.9437	0.9665	0.7409	0.7363	0.9709	–	0.8462	0.7003
Vega [30]	0.9483	0.9671	0.7019	–	–	0.6616	0.8240	0.6400
You [31]	0.9497	0.9750	0.7260	–	–	–	0.8414	–
Al-Diri [32]	–	0.9681	0.7521	–	–	–	0.8533	–
Orlando [15]	–	0.9738	0.7680	0.7740	–	0.7644	0.8628	0.7417
2nd	0.9352	0.9387	0.8951	0.6425	0.9884	0.7401	0.9142	0.7225
Work [20]	0.9364	0.9627	0.7046	0.6984	0.9665	0.6904	0.8202	0.6620
Proposed	0.9477	0.9733	0.7202	0.7566	0.9689	0.7260	0.8312	0.7045

defined as:

$$ACC = (TN + FP)/(TN + TP + FN + FP) \tag{11}$$

$$SP = TN/(TN + FP) \tag{12}$$

$$SE = TP/(TP + FN) \tag{13}$$

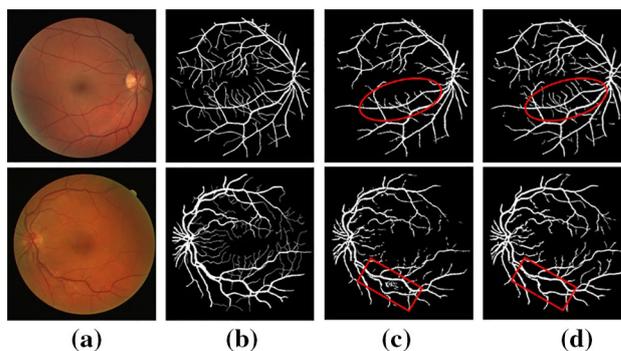
$$PPV = TP/(TP + FP) \tag{14}$$

$$NPV = TN/(TN + FN) \tag{15}$$

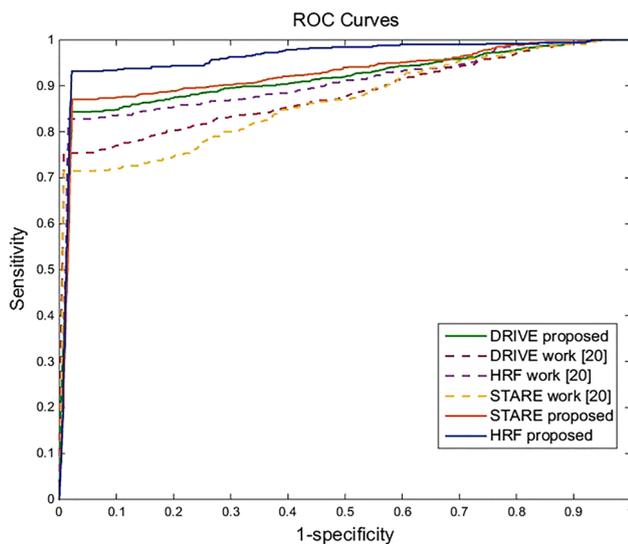
$$F1 = (2 \times PPV \times SE)/(PPV + SE) \tag{16}$$

$$G = \sqrt{SE \times SP} \tag{17}$$

$$MCC = (TP/M - S \times P)/(P \times S \times (1 - S) \times (1 - P)) \tag{18}$$



**Fig. 7** Segmentation comparisons, **a** original images, **b** manual annotations, **c** segmentation results in paper [20], **d** segmentations using the proposed method



**Fig. 8** ROC comparisons between work [20] and the proposed method on HRF, DRIVE, and STARE databases

### 3.3 Segmentation results

Figures 4, 5 and 6 show partial segmentation results on HRF, DRIVE and STARE databases. As can be observed, our method can extract abundant vascular branches with different thickness by adopting multiple features for the discriminative dictionary learning. However, for pathological cases in STARE, our method may incur misclassified non-vessel structures in hemorrhage areas as presented in Fig. 6. Also, we can notice that vessels segmented in HRF obtains more thin vessels than STARE and our method may have slightly block effect.

### 3.4 Comparison with other methods

We have conducted comparisons with our previous work [20] and other methods. Tables 1, 2 and 3 show segmentation comparisons on HRF, DRIVE and STARE respectively (2nd

means the second annotation). Here we can obtain the following observations: (1) For DRIVE, despite our method shows a lower SP compared with Fraz's methods [7,29], but it still has a larger SE. As mentioned before, SE represents the capability for distinguishing vascular details. (2) For STARE, our method has relatively higher PPV and SP. Since SP measures the capability for distinguishing non-vessels, it means that our method obtains less false positive pixels. (3) For HRF, since limited methods are reported in the literature, our method gives higher metric values than other methods except for SP. (4) Compared with our previous work, metric values are significantly improved on all databases. Particularly, the improvement in SE is larger than that in SP and PPV, which shows that more vessel details are correctly detected as shown in Fig. 7. The reason may be that only a single feature map is used for training in our previous work, but in the proposed method, we adopt multiple feature maps and add vessel thickness labels for training with LC-KSVD instead of KSVD. Figure 8 shows the ROC comparisons. (5) The deep learning method [20] shows better performance on STARE and DRIVE databases than our method.

### 3.5 Cross-database validation

To further analyze the performance, we carry out cross-database validation for assessment. We use dictionary learned from one dataset (e.g., DRIVE, HRF or STARE) to test performance on another. Results are presented in Table 4. As can be seen, sensitivity in all databases is decreased when using other training databases for training. For example, sensitivity on HRF is decreased to 0.7626 when trained on DRIVE, and 0.4584 when trained on STARE. The reason is that HRF has thinner vessels compared with DRIVE, and STARE contains more pathological and low-quality images. Also, we select 10 images from each database to construct a hybrid training set with 30 images using leave-one-out strategy. As can be observed, the performance of hybrid training is acceptable. This phenomenon indicates that selecting suitable training

**Table 4** Performance of cross-database validation

Database	ACC	SP	SE	PPV	NPV	F1	G	MCC
HRF (trained on DRIVE)	0.9505	0.9692	0.7626	0.7119	0.9762	0.7322	0.8591	0.7084
STARE (trained on DRIVE)	0.9454	0.9758	0.6654	0.7736	0.9643	0.6904	0.7930	0.6769
DRIVE (trained on HRF)	0.9423	0.9822	0.6769	0.8492	0.9542	0.7499	0.8142	0.7263
STARE (trained on HRF)	0.9488	0.9818	0.6455	0.8170	0.9624	0.6935	0.7817	0.6866
HRF (trained on STARE)	0.9367	0.9772	0.4584	0.6552	0.9554	0.5225	0.6645	0.5088
DRIVE (trained on STARE)	0.9354	0.9870	0.3240	0.6773	0.9455	0.4231	0.5523	0.4309
HRF (hybrid training)	0.9492	0.9727	0.7039	0.7163	0.9717	0.7059	0.8264	0.6806
DRIVE (hybrid training)	0.9405	0.9730	0.7188	0.7973	0.9595	0.7527	0.8353	0.7223
STARE (hybrid training)	0.9472	0.9799	0.6535	0.7985	0.9624	0.6991	0.7905	0.6855

samples is essential. In the future, we expect to construct a larger training dataset that can achieve stable performances on all databases.

## 4 Conclusions

In this paper, we have presented a new retinal blood vessel segmentation method that segments vessel structures via dictionary learning. Rather than using conventional dictionary learning framework, our approach adopts multiple features to learn the discriminative dictionary. Our contributions are mainly in two aspects. First, we use LC-KSVD framework that adds a vessel thickness label term to supervise dictionary learning. Second, we extract multiple feature maps for feature representation. Experiments on benchmark databases demonstrate that our method can yield good segmentation results. In the future, we plan to design better sample selection and feature representation methods for training. We also expect to apply our method on other images to improve the cross-database capability.

**Acknowledgements** This work was supported by the Natural Science Foundation of China (Grant 61622109), the Zhejiang Natural Science Foundation of China (Grant R18F010008), and the Natural Science Foundation of Ningbo (2017A610112). It was also sponsored by K.C. Wong Magna Fund in Ningbo University.

## References

1. Abràmoff, M.D., Garvin, M.K., Sonka, M.: Retinal imaging and image analysis. *IEEE Rev. Biomed. Eng.* **3**, 169–208 (2010)
2. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A.: Blood vessel segmentation methodologies in retinal images: a survey. *Comput. Methods Prog. Biol.* **108**, 407–433 (2012)
3. Annunziata, R., Garzelli, A., Ballerini, L.: Leveraging multiscale Hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation. *IEEE J. Biomed. Health Inf.* **20**, 1129–1138 (2016)
4. Yu, H., Barriga, S., Agurto, C., Zamora, G., Bauman, W., Soliz, P.: Fast vessel segmentation in retinal images using multiscale enhancement and second-order local entropy. *Proc. SPIE* **8315**, 83151B-1–83151B-12 (2012)
5. Azzopardi, G., Strisciuglio, N., Vento, M.: Trainable COSFIRE filters for vessel delineation with application to retinal images. *Med. Image Anal.* **19**, 46–57 (2015)
6. Zhao, Y., Zhao, J., Yang, J., Liu, Y.: Saliency driven vasculature segmentation with infinite perimeter active contour model. *Neurocomputing* **259**, 201–209 (2017)
7. Fraz, M.M., Basit, A., Barman, S.A.: Application of morphological bit planes in retinal blood vessel extraction. *IET Image Process.* **26**, 373–383 (2013)
8. Zhao, Y., Zheng, Y., Liu, Y., Zhao, Y., Luo, L., Yang, S.: Automatic 2D/3D vessel enhancement in multiple modality images using a weighted symmetry filter. *IEEE Trans. Med. Imaging* **37**, 438–450 (2018)
9. Zhang, J., Dashtbozorg, B., Bekkers, E., Pluim, J.P.W., Duits, R.: Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores. *IEEE Trans. Med. Imaging* **35**, 2631–2644 (2016)
10. Soares, J., Leandro, J., Cesar, R., Jelinek, H.F., Cree, M.J.: Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification. *IEEE Trans. Med. Imaging* **25**, 1214–1222 (2006)
11. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B.: An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Trans. Med. Imaging* **59**, 2538–2548 (2012)
12. Aslani, S., Sarnel, H.: A new supervised retinal vessel segmentation method based on robust hybrid features. *Biomed. Signal Process. Control* **30**, 1–12 (2016)
13. Roychowdhury, S., Koozekanani, D.D., Parhi, K.K.: Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. *IEEE J. Biomed. Health Inf.* **19**, 1118–1128 (2015)
14. Marin, D., Aquino, A., Gegundez-Arias, M.E., Bravo, J.M.: 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)
15. Orlando, J.I., Prokofyeva, E., Blaschko, M.B.: A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE Trans. Med. Imaging* **64**, 16–27 (2017)
16. Li, Q., Feng, B., Xie, L., Liang, P.: A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Trans. Med. Imaging* **35**, 109–118 (2016)
17. Fu, H., Xu, Y., Wong, D., Liu, J.: Deep vessel: retinal vessel segmentation via deep learning and conditional random field. *Proc. MICCAI* **9901**, 132–139 (2016)
18. Yan, Z., Yang, X., Cheng, K.: Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation. *IEEE Trans. Med. Imaging* **65**, 1912–1923 (2018)
19. Shao, F., Yang, Y., Jiang, Q.: Automated quality assessment of fundus images via analysis of illumination, naturalness and structure. *IEEE Access.* **6**, 806–817 (2018)
20. Yang, Y., Shao, F., Fu, Z., Fu, R.: Blood vessel segmentation of fundus images via cross-modality dictionary learning. *Appl. Opt.* **57**, 7287–7295 (2018)
21. Guen, V.L.: Cartoon + texture image decomposition by the TV-L1 model. *IPOL* **4**, 204–219 (2014)
22. Frangi, A., Niessen, W., Vincken, K.: Multiscale vessel enhancement filtering. *Med. Image Comput. Comput. Assist. Interv.* **1496**, 130–137 (1998)
23. Wang, L., Xiao, L., Liu, H., Wei, Z.: Variational Bayesian method for Retinex. *IEEE Trans. Image Proces.* **23**, 3381–3396 (2014)
24. Zhou, M., Jin, K., Wang, S., Ye, J., Qian, D.: Color retinal image enhancement based on luminosity and contrast adjustment. *IEEE Trans. Biomed. Eng.* **65**, 521–527 (2018)
25. Jiang, Z., Lin, Z., Davis, L.: Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 2651–2664 (2013)
26. Odstrcilik, J., Kolar, R., Budai, A., Hornegger, J.: Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Process* **7**, 373–383 (2013)
27. Staal, J., Abramoff, M., Niemeijer, M., Viergever, M.: Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **23**, 501–509 (2004)
28. Hoover, A., Kouznetsova, V., Goldbaum, M.: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. Med. Imaging* **19**, 203–210 (2000)

29. Fraz, M.M., Remagnino, P., Hoppe, A., Uyyanonvara, B.: Retinal vessel extraction using first-order derivative of Gaussian and morphological processing. *Adv. Vis. Comput.* **6938**, 410–420 (2011)
30. Vega, R., Sanchez-Ante, G., Falcon-Morales, L.: Retinal vessel extraction using lattice neural networks with dendritic processing. *Comput. Biol. Med.* **58**, 20–30 (2015)
31. You, X., Peng, Q., Yuan, Y.: Segmentation of retinal blood vessels using radial projection and semi-supervised approach. *Pattern Recogn.* **44**, 2314–2324 (2011)
32. Al-Diri, B., Hunter, A., Steel, D.: An active contour model for segmenting and measuring retinal vessels. *IEEE Trans. Med. Imaging* **28**, 1488–1497 (2009)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.