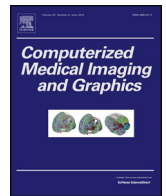




Contents lists available at SciVerse ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag



Automated characterization of blood vessels as arteries and veins in retinal images

Qazaleh Mirsharif^{a,*}, Farshad Tajeripour^{a,2}, Hamidreza Pourreza^{b,3}

^a Department of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

^b Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE INFO

Article history:

Received 26 October 2012

Received in revised form 28 April 2013

Accepted 17 June 2013

Keywords:

Artery/vein classification

Blood vessels

Retinal images

Retinex image enhancement

Feature extraction

Structural features

ABSTRACT

In recent years researchers have found that alternations in arterial or venular tree of the retinal vasculature are associated with several public health problems such as diabetic retinopathy which is also the leading cause of blindness in the world. A prerequisite for automated assessment of subtle changes in arteries and veins, is to accurately separate those vessels from each other. This is a difficult task due to high similarity between arteries and veins in addition to variation of color and non-uniform illumination inter and intra retinal images. In this paper a novel structural and automated method is presented for artery/vein classification of blood vessels in retinal images. The proposed method consists of three main steps. In the first step, several image enhancement techniques are employed to improve the images. Then a specific feature extraction process is applied to separate major arteries from veins. Indeed, vessels are divided to smaller segments and feature extraction and vessel classification are applied to each small vessel segment instead of each vessel point. Finally, a post processing step is added to improve the results obtained from the previous step using structural characteristics of the retinal vascular network. In the last stage, vessel features at intersection and bifurcation points are processed for detection of arterial and venular sub trees. Ultimately vessel labels are revised by publishing the dominant label through each identified connected tree of arteries or veins. Evaluation of the proposed approach against two different datasets of retinal images including DRIVE database demonstrates the good performance and robustness of the method. The proposed method may be used for determination of arteriolar to venular diameter ratio in retinal images. Also the proposed method potentially allows for further investigation of labels of thinner arteries and veins which might be found by tracing them back to the major vessels.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Retinal image analysis has been an active field of research in the scope of medical image processing for years since it has been realized that alternations in retinal vascular network are associated with several cardiovascular disorders. Recently, ophthalmologists have assigned the small changes in the caliber of arteries and veins

to several micro-vascular diseases such as diabetic retinopathy. Diabetic retinopathy is an increasingly growing public health problem and the leading cause of blindness in the world. Also other diseases such as atherosclerosis and hypertension affect arteries and veins differently leading to an abnormal artery to vein width ratio (AVR) in retinal vasculature [1–4]. Ophthalmologists suggest that cardiovascular diseases may affect the length of vessels, increase their curvature and tortuosity, or modify the appearance of vessels. Therefore investigation of changes in arteries and veins might lead to significant foundations and deserves to be explored.

Manual assessment of subtle changes in arteries and veins is a cumbersome task and requires highly trained personnel. As a result, developing computer algorithms for this purpose are of paramount importance in assisting doctors for early identification and timely treatment of the associated diseases. An important pre-requisite for automated measurement of small changes in each type of vessels, is separation of arteries from veins. This task must be performed precisely in order to detect early signs of the mentioned diseases.

Automated classification of retinal blood vessels into arteries and veins is a difficult task and has been less explored in the

* Corresponding author. Tel.: +98 9153197716.

E-mail addresses: Qazaleh.mirsharif@gmail.com, mirsharif@shirazu.ac.ir (Q. Mirsharif), tajeri@shirazu.ac.ir (F. Tajeripour), hpourreza@um.ac.ir, hpourreza@ieee.org (H. Pourreza).

URLs: <http://pasargad.cse.shirazu.ac.ir/mirsharif/> (Q. Mirsharif), <http://research.shirazu.ac.ir/Faculty/Detail.asp?ID=871> (F. Tajeripour), <http://hpourreza.profcms.um.ac.ir/> (H. Pourreza).

¹ Address: Rezashahr st, South rudaki 17, number 58, Mashhad, Khorasan razavi 9175846463, Iran.

² Address: Molasadra st, Computer Engineering Department at Shiraz University, Shiraz, Iran. Tel.: +98 7116133596.

³ Room: 407. Tel.: +98 5118805025; fax: +98 5118763306.

literature. Arteries and veins are very similar in appearance. In addition, the curved shape of retina and inappropriate illumination conditions during image acquisition process might lead to non uniform illumination through retinal images. Also biological characteristics, in this case changes in color of retina from person to person, raise another problem [5,6].

There have been a number of methods reported in the literature for artery/vein classification of retinal blood vessels which fall into two categories: automated and semi-automated methods. The automated methods are based on feature extraction from major vessels. In those methods first, vascular network is separated from image background and other structures in the image. Then the centerline pixels or skeleton of vessels are extracted from the segmented vascular tree. For each centerline pixel, various features are calculated and finally each centerline pixel is assigned an artery or a vein label by classifier. In a semi-automated approach, first, initial points of main vessels are labeled by ophthalmologists as arteries or veins. Then those labels are propagated toward smaller vessels using the structural characteristics and connectivity information of the vascular tree. According to structural features of vascular tree, an artery or vein never crosses a vessel of the same type which means in a cross over point one vessel is artery and the other one is vein. Also three vessel segments connected to each other through a bifurcation point are of the same type of vessel.

The research work carried by Grisan and Ruggeri in 2003 [7] was among first attempts to separate arteries from veins automatically. The main idea behind their method was the symmetric division of the optical disk (OD) based retinal images into four partitions, feature extraction from vessels and applying fuzzy clustering classification to vessels in each partition independently. The results in each region were then combined and the total error rate of 12.4% for artery/vein classification in 24 images was reported. In another paper, Li et al. [8], adjusted a piece-wise Gaussian model to capture the central reflex feature of blood vessels which is more obvious in arteries than veins and Jelinek et al. [9], tested different classifiers and features in RGB and HSL color spaces for discrimination of arteries from veins. Also Konderman et al. [10] 2007 used features based on vessel profile at each centerline pixel as well as intensity values of pixels around those centerline points for classification of arteries and veins. They achieved 96.32% accuracy for classification of arteries and veins in 4 images. These methods are not comparable since they have been evaluated on different datasets most of which contain a few images.

In 2007 Rothaus et al. [11], proposed a semi-automated method for separation of artery map from veins. They assumed the vascular tree as a graph where intersection and bifurcation points are considered as vertices of the graph. The vessel segments connected at vertices were treated as edges of the graph. Utilizing an optimization problem they attempted to publish the manually labeled initial vessel points through the graph leaving the minimum conflicts possible using structural features of vessel tree. They have reported their results as the number of unresolved conflicts in vessel map. Niemeijer et al. [12], extracted features such as pixel intensity and average intensity values on profile of vessels in HSL color space. They have also exploit steerable Gaussian features from profiles of vessels at different angles in Green plane and classified the major vessels using kNN classifier. They have evaluated their method on publicly available database of DRIVE and achieved 88% area under ROC curve. V'azquez et al. [13], combined several feature vectors and clustering algorithms in regions of different radii from optical disk. They achieved the best results by dividing the image into four partitions and applying feature extraction to each partition separately repeating this process by rotating the image. Combining the results in the resulted overlapping regions they reported classification accuracy of 86.34% on VICAVER dataset which consists of 58 OD based retinal images. V'azquez et al. [14], raised this accuracy

up to 91% by applying multi-scale retinex image enhancement to the images before feature extraction step.

Accurate classification of arteries from veins has been studied mainly for measurement of arteriolar-to-venular diameter ratio (AVR) [15–18] so far and is based on feature extraction from major vessels. The process of AVR estimation follows segmentation of vessels in a region of interest for measuring AVR [19,20] separation of major arteries from veins in the so-called region, and measuring the mean diameter of vessels in each class for calculation of final arteriolar-to-venular diameter ratio. However, extending those methods to a broader range of vessels is of high interest to ophthalmologists since will allow them to study the properties of each type of vessels independently and analysis the probable effects of micro vascular complications such as changes in curvature, branching pattern or appearance of vessels. For this purpose, structural characteristics of vessel network are exploited in order to publish the labels to smaller vessels that are hardly recognizable using their color features. Methods which use structural features are dependent on the quality of vessel segmentation and how much correctly labeled the major vessels are. Lack of information for finding connectivity of vessels at cross over and bifurcation points as well as publishing a wrong label may decrease the recognition rate considerably. Methods such as [21] use the manually labeled vessels as start points for spreading the labels. The method also requires user interference for revising the incorrectly labeled artery or vein trees found by their algorithm. To develop an automated method, initial points of vessels must be labeled using feature extraction. This process however, increases the probability of publishing a wrong label comparing to semi-automated method and thus requires the main vessels from which the thinner ones are traced to be classified with a high accuracy.

In this paper we present an automated method for classification of relatively major retinal vessels into arteries and veins. We intend to develop an automated method for separation of all vessels from which the secondary or small vessels are traced while we maintain the recognition rate of main vessel in the region of interest for AVR measurement comparable with previous methods. The final output of the proposed method is two binary images of artery and vein maps which can be used as start maps for spreading labels of major vessels toward small ones. Also the method may be used for determination of AVR.

The major contributions of the paper are as follow:

1. The proposed method in this paper achieves high classification rate without increasing the training samples or adding many features. By dividing the vessels many redundant sample points are removed.
2. Combination of image enhancement techniques such as histogram matching, CLAHE and MSRCR improves the results significantly.
3. We present an automated method which combines color features with structural features of vessel networks that leads to better results.
4. Also in this paper a new region separation scheme is proposed for improving the clustering classification of retinal blood vessels.

This paper is organized as follows: Section 2 is devoted to description of two datasets of retinal images used in this paper including publicly available dataset of DRIVE. In Section 3 the proposed method is explained which consists of three main steps including image enhancement, classification of vessels into arteries and veins and a post processing step for improving the classification rate obtained from previous step. In Section 4 experimental results are discussed in detail. Ultimately Section 5 concludes the research.

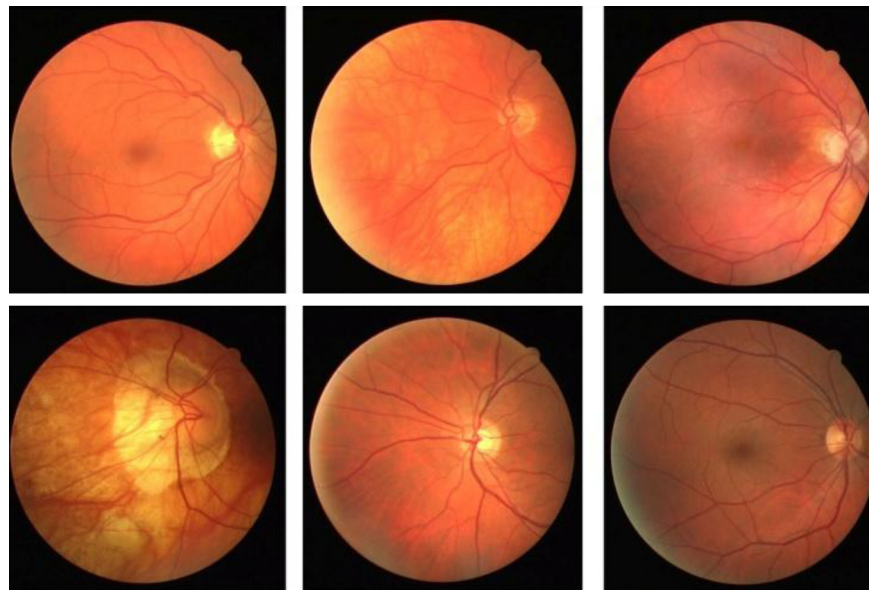


Fig. 1. Sample images of DRIVE database. Note the high color and illumination variation inter and intra images.

2. Dataset

In this paper two different datasets of retinal images have been used for evaluation of the proposed methods which will be explained in this section. The first dataset is DRIVE which is publicly available and has been widely used for evaluation of algorithms developed to deal with different problems in the area of retinal image processing such as retinal vessel extraction. We also use a second dataset which has been provided in Khatam-Al-Anbia eye hospital of Mashhad, Iran to evaluate the robustness and performance of our proposed method on different images.

2.1. DRIVE database

DRIVE database [35] was introduced in 2004 and is publicly available for evaluation of research works on retinal vessel segmentation. This database includes 20 train and 20 test images which are selected randomly from 400 photographs taken from diabetic subjects. 7 of those images show early sign of diabetic retinopathy. For each of those 40 images, there is a binary map of vessels that is segmented manually and can be used as reference standard. To compare the result of automated vessel classification, our ophthalmologist separated relatively major vessels using blue color for veins and red for arteries. These images are used as gold standard in this paper. Some sample image from DRIVE database can be seen in Fig. 1.

2.2. Retinal images of Khatam

Khatam dataset has been obtained in Khatam-Al-Anbia eye hospital of Mashhad, Iran and currently include 13 retinal images with the resolution of 2592×3872 . For every image two separate binary maps of arteries and veins have been provided under the supervision of an ophthalmologist for evaluation of results of the proposed method. Also the location of optical disk in each image is determined manually. This image set is used to evaluate the robustness of our method. See Fig. 2.

3. The proposed method

In this section, the proposed method is presented which consists of three steps: Image enhancement, vessel classification and post-processing. Each step of these stages will be explained in detail.

The general approach an ophthalmologist utilizes for separation of arteries from veins is as follows: First the optical disk in image is detected. Optical disk is a bright approximately oval shape area in retina which acts as an entrance region for vessels. Thus the initial points of major vessels can be found around optical disk and traced outwards to the thinner vessels. Major vessels in the region around the (O.D.) can be discriminated based on features such as color, illumination, central reflex and width of vessels. Generally, arteries are brighter and thinner than veins and the central reflex at the inner part of vessels is more obvious in arteries since they carry blood rich in oxygen.

The ophthalmologists then track the major vessels toward thinner ones. Tracking the vessels might be based on both color and width of vessels and structural features at cross over and bifurcation points in the vascular network. As the vessels become smaller, tracking become dependent mostly on structural features at cross over and bifurcation points. Cross over points are the points where two vessels of different types cross each other. In bifurcation points a vessel branches out into two thinner vessels. The structural characteristics of vessels are used as rules explained below:

- Arteries never cross arteries and the same is true for veins. Thus in any cross-over points where two vessels cross each other, one of them is artery and another one is vein.
- The retinal blood vasculature follows the structure of a binary tree. Therefore, a vessel branches into two vessels of the same type which means, in every bifurcation point, three vessel segments connected to each other are from the same class of vessel.

Fig. 3 illustrates the semi automated process of artery/vein classification. The semi-automated approach as was explained earlier is based on tracking initial vessel points that are manually labeled by ophthalmologists toward thinner vessels. The first idea that comes to mind for developing an automated method for the artery/vein

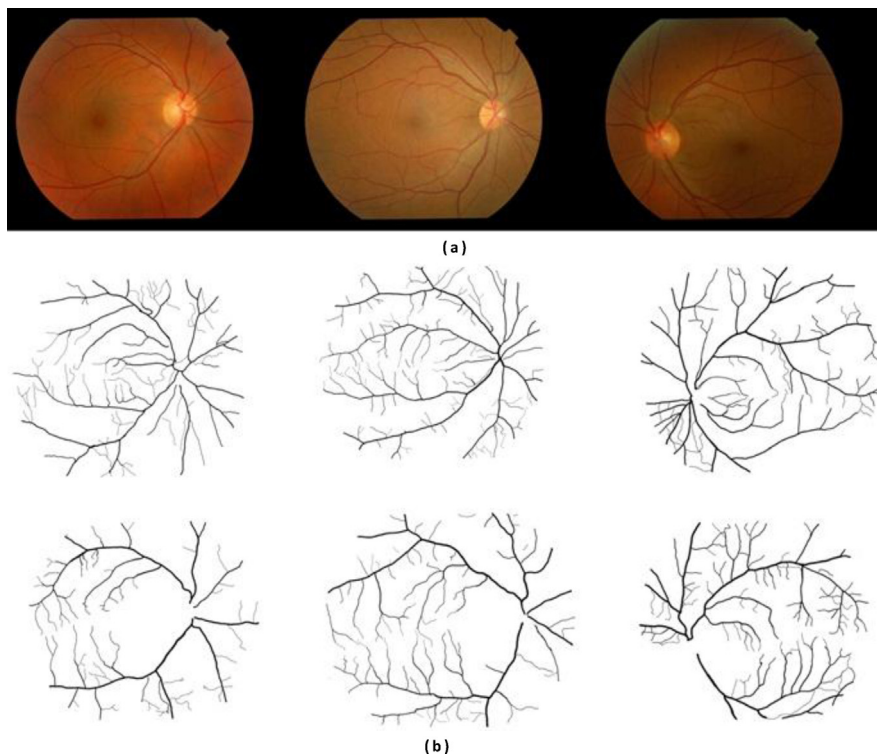


Fig. 2. (a) Sample images of Khatam database. Note the high color and illumination variation inter and intra images. (b) The second and third rows respectively show the artery and vein maps separated manually by ophthalmologists.

classification of retinal blood vessels may be to classify the major vessels near optical disk and propagate their labels through the whole vessel tree using color and structural features. However, if the initial points are labeled incorrectly by the automated feature based method, a wrong label will be published through the vessels. Therefore, we attempt to classify the relatively major vessels all over the retinal images and connect the vessels of the same type through finding sub trees of arteries and veins. Indeed after the initial classification of vessels using their color features, the dominant label in each vessel sub tree is propagated through vessel points of that sub tree.

In the proposed method, first the contrast and quality of the images are improved. Then vessel network is extracted from the

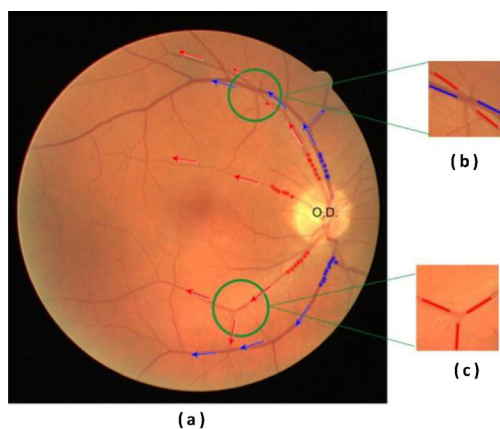


Fig. 3. (a) A sample RGB retinal image for which the semi-automated tracking based artery/vein classification is shown. As can be seen the main vessels near the optical disk are labeled and tracked toward the thinner ones. (b) Two vessels cross each other and make four vessel segments two of which are certainly veins and the other two segments are arteries. (c) A bifurcation point where three vessel segments are connected. All of the three vessels are of the same type.

image background and major vessels are classified into arteries and veins. Finally connected sub trees of vessels are found using structural features and the dominant label in each sub tree of vessels is published through the entire tree. Fig. 4 shows stages of the proposed method.

Next the first stage of the proposed algorithm which is image enhancement will be explained followed by feature extraction and post processing steps.

3.1. Image enhancement

Before classification of blood vessels into arteries and veins we attempt to improve the retinal images by evaluation of several image enhancement techniques. Image enhancement is an important step in our method since the contrast between arteries and veins in the retinal images are initially very low. Also the variation of color and illumination inter and intra images significantly affect the recognition rate of the automated methods [5–7].

For color normalization through images a histogram matching algorithm is applied. Histogram matching algorithm takes two images including the source image A and the reference image R and returns image B as the output. Using the histogram matching algorithm, image A is transformed so that the histogram of



Fig. 4. Stages of the proposed method.

resulted image B approximately matches the histogram of the reference image R . The reference image is chosen subjectively by ophthalmologist as an image which is better in quality and contrast comparing to other images in the training set. In this paper the reference image has been selected from DRIVE database and is used for histogram matching in both DRIVE and Khatam datasets. The histogram matching method is applied to each individual RGB image. For each RGB channel histogram matching is employed independently which means that the histograms of red, green and blue channels are matched respectively with histograms of red, green and blue channels of the reference RGB image.

After color image normalization, three different techniques including generalized histogram equalization (GHE), piece-wise histogram equalization (PHE) and adaptive histogram equalization (AHE) are applied to the images for improving their contrast. GHE is widely used in image processing applications for image enhancement. GHE changes the intensity distribution in order to make it uniform to its utmost degree.

One of the variations of histogram equalization is piece wise histogram equalization (PHE). These methods are usually more effective than simple histogram equalization and faster than local methods. In such methods, histogram of image is divided to sub histograms based on different criterions such as local minima points of the histogram. Then each sub histogram is equalized separately.

Wu et al. [22] proposed weighting mean-separated sub histogram equalization method for contrast enhancement. They used Eq. (1) for determining each sub histogram points:

$$X_t = \frac{\sum_{l=a}^b l \times CDF(l)}{\sum_{l=a}^b CDF(l)} \quad (1)$$

where t is the recursion level, $[a\ b]$ is each sub histogram level initialized at $[0\ 255]$, l is gray level and CDF represents the cumulative density function for histogram of image.

Afterwards each sub histogram is equalized using the weighted transform function below:

$$T_k(G_k) = X_k + (X_{k+1} - X_k) \times (CDF_k(G_k)) \quad (2)$$

G_k represents the gray level for the respective sub histogram $[X_k\ X_{k+1}]$.

We also examine a fully automated piece-wise linear histogram equalization method introduced in [23] for image enhancement in color images. In their method, major picks of the histogram are used as midpoints for histogram segmentation. First the histogram of image in each channel is obtained. Afterwards a smoothing function is used to omit the frequent non-important picks of the histogram. The input points are determined by the intensity values of the respective picks and the output ones are indicated using Eq. (3). Finally each sub histogram is equalized piece wisely using the new transform function defined by Eq. (4)

$$Y_k = \sum_{x=0}^{v_k} Pr(x) \times 255 \quad (3)$$

$$T_{k-1}(x) = \frac{Y_k - Y_{k-1}}{X_k - X_{k-1}} \times (X - X_{k-1}) + Y_{k-1} \quad (4)$$

where $[X_k\ X_{k-1}]$ represent interval for each sub histogram. Y_k represents the output for candidate midpoints and Pr is the probability density function.

Their results are better comparing to GHE. Finally CLAHE is applied which leads to better results comparing to GHE and PHE. CLAHE operates on small regions in the image, called *tiles*, rather than the entire image. Each tile's contrast is enhanced, so that the histogram of the output region approximately matches the histogram specified by the 'Distribution' parameter. The neighboring

tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image. As can be seen in Fig. 5 the output images from CLAHE technique look better. Also objective experiments confirm the superiority of the CLAHE comparing to the other methods. For further information refer to [24].

Although the contrast between arteries and vein increases, non uniform illumination still remains a problem in the retinal images. To lessen the effect of inhomogenous illumination Multi-scale retinex with color restoration (MSRCR) image enhancement technique is employed [25,26]. Multi-scale retinex (MSR) method has gained special attention in recent years for shadow removing and to overcome the problem of non-uniformity of illumination in image. The idea behind Retinex theory is to approximate the illumination in image using Gaussian function and subtracting the estimation from the original image. The output is an image invariant to illumination. The single scale retinex output is obtained using:

$$R_i = \log I_i(x, y) - \log[F(x, y) * I_i(x, y)] \quad (5)$$

$$F(x, y) = Ke^{-\frac{x^2+y^2}{\sigma^2}} \quad (6)$$

where R_i represents the output of single scale retinex, I is the input image, F represents the surround function, sigma is the scale for controlling the surround function and i is the respective index for each RGB plane. The multi-scale version is the summation of single scale retinex (SSR) output in multiple desired scales as follow:

$$R_m = \sum_{n=1}^N W_n R_{n_i} \quad (7)$$

where W_n is the respective weight for each SSR output and N is the number of scales. As was discussed by Jobson et al. [25], usually three scales (small, medium and large) are enough for the most images. MSRCR combines the dynamic range compression of the small-scale retinex and the tonal rendition of the large scale retinex with a universally applied color restoration. Indeed, color restoration is added to preserve a reasonable degree of color constancy in the images.

$$MSRCR_i = G[C_i(\log I_i - \log[I_i * F_n]) + b] \quad (8)$$

where G and b are the final gain and offset values and:

$$C_i = \beta \left\{ \log[\alpha I_i] - \log \left[\sum_{i=1}^s I_i \right] \right\} \quad (9)$$

where β is a gain constant and α controls the strength of the non-linearity. V'azquez et al. [14], previously discussed the influence of single scale and multi-scale retinex on retinal images of VICAVR dataset and the result demonstrated that the images were best improved using a multi scale retinex with three scales including small, middle and large scales. In our method, The MSRCR approach is applied to each channel of RGB color space for every individual retinal image in DRIVE and Khatam-Al-Anbia datasets. The parameters are set using an evaluation set of data so that the color information is preserved while the illumination condition improves in the images.

Fig. 5 demonstrates resulted images after applying CLAHE and MSRCR on two mainly used planes for classification of vessels including red and green. Blue channel usually contains much noise and is removed from the process of feature extraction. Later objective comparisons also demonstrate that applying MSRCR after CLAHE improves the classification rate of vessels substantially in red and green planes [24].

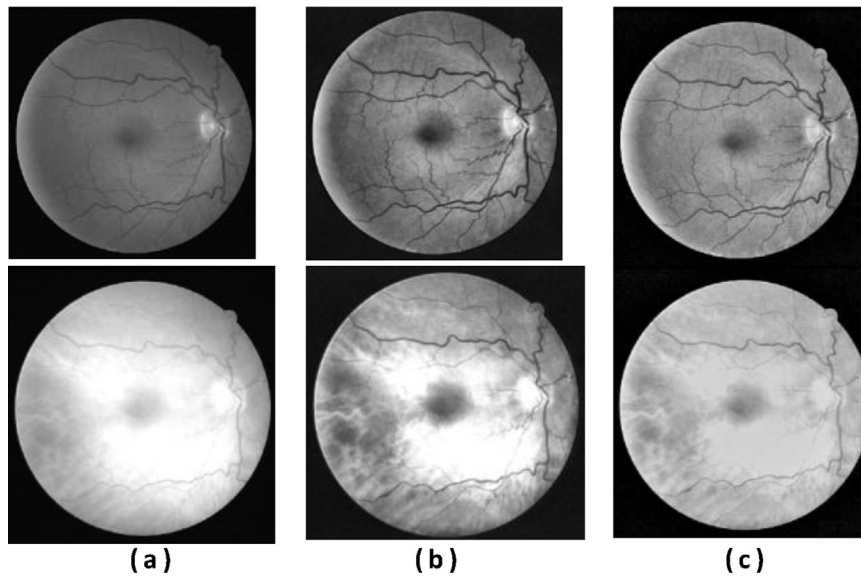


Fig. 5. (a) First column shows sample images in Red and Green channels of RGB color space. (b) Second column shows respective images after enhancement using CLAHE and before applying MSRRCR. (c) Third column shows the resulted images after applying CLAHE and MSRRCR image enhancement. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.2. Feature extraction and vessel classification

There have been a number of methods in the literature proposed for artery/vein classification of blood vessels. The automated approaches are based on feature extraction from major vessels. Particularly most of those methods concentrate on classification of vessels in the region of interest for AVR measurement.

For classification of vessels first vascular structure must be extracted from image as a binary image which indicates the location of vessel pixels. There have been many approaches presented for vessel extraction from retinal images [27–30]. In this paper we have used the method proposed by Soares et al. which uses Gabor wavelets for feature extraction [31].

After vessel extraction small vessel thinner than three pixels are removed from the vessel network using morphological structures and a thinning algorithm [32,33] is applied to the rest of vessels for extraction of their centerline points. The thinning algorithm removes the outer pixels of the vessels until a structure of one pixel thickness is obtained. The thinning method used in this paper maintains the connectivity of the vessels. See Fig. 6. The skeleton of vessels is used for feature extraction and tracking of the vessels through the vessel tree.

After extraction of centerline pixels of vessels, bifurcation and cross-over points are discarded from vessel skeleton. Cross over and bifurcation points are the pixels in skeleton for which there are more than two adjacent pixels in the skeleton. Indeed, they

indicate where two vessels pass each other or a vessel branches into two thinner vessels. The output of this stage contains a binary image of vessel segments.

Most of the previous studies discriminate vessels based on pixel classification which means for each centerline pixel features are extracted and assigned to that pixel. The extracted features are usually obtained by taking the average or variance of intensity values of pixels located on profile of vessels or neighbor pixels around the centerline ones. Profile is a one pixel thick line perpendicular to the vessel direction. See Fig. 7. In addition to profile based features from RGB planes of color images, pixel intensity features from HSL color space have been employed for artery/vein classification [7]. Some methods used the Gaussian derivatives of pixel intensities at different angles that are suitable for capturing the central reflex property of blood vessels [12]. Arteries and veins can be discriminated the best using the red plane since veins have dark red color and arteries are bright red. However, since we have great loss of information for some images in the red channel and this channel is very saturated for some images and thus has less space to improve, green plane must be added for feature extraction. We also investigate the HSL and LAB color spaces since the contrast between arteries and veins in some green images is very low.

All the pixels in each vessel segments are connected and of the same type, and usually features do not change much for consecutive pixels in a close neighborhood. Therefore, we divide each vessel segment into smaller vessels and extract our features for new sub

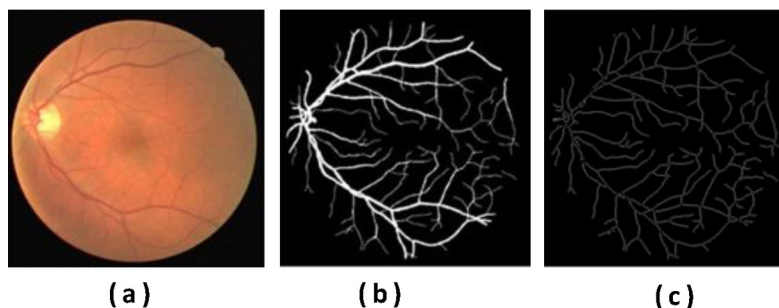


Fig. 6. (a) A sample retinal image. (b) Binary map of vascular tree for the image in (a). (c) Skeleton of vascular tree for the respective image.

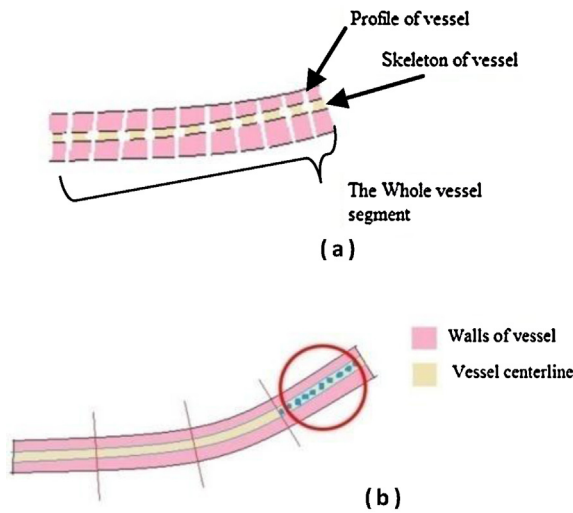


Fig. 7. (a) In previous works features were extracted from pixels on profile of vessel for every centerline pixel. The label assigned to the majority of centerline pixels by the classifier was then propagated through the whole vessel segment. (b) In this work each vessel segment is divided to smaller sub vessels and features are extracted over pixel in the skeleton of sub vessel or in the whole sub vessel inside the circular window.

vessels instead of each centerline pixel. See Fig. 7. Each vessel segment consists of an inner part and outer part or walls of vessels. The color and illumination in the inner part and walls of vessels are different for arteries and veins. Since arteries carry blood rich in oxygen their inner part is brighter than their walls comparing to veins which means the central reflex feature is more obvious in arteries.

After dividing the vessel, for each new sub vessel features such as mean and variance of intensity values for the centerline pixels as well as the pixels of the entire sub vessel are extracted and assigned to that sub vessel. Also difference of the average value of pixel intensities over centerline pixels and walls of vessels are obtained in each planes of image in RGB, HSL and LAB color spaces and added to the feature vector. Using mean and variance of pixel intensities for a broader range of pixels comparing to profile or pixel based features decrease the effect of noise and outliers in the classification process. For calculation of mean and variance values on inner points of vessels, skeleton of vessels is exploited since the skeleton has much overlap with the centerline pixels of the vessels. For measurement of other features a window of size 12*12 is located on the middle point of each sub vessel and the mean and variance are calculated for the pixels of vessels inside the window [34]. See Fig. 7.

Finally, a forward feature selection method is used for selection of the most discriminant features for training the classifier. Each feature is added to the feature vector. If the accuracy for the evaluation set of images increases, that feature is considered for the final feature vector. Table 1 shows the best features that have been chosen by feature selection process and have been used for the final classification of vessels. As is seen in Table 1, most selected features are extracted from red and green channels which means arteries are recognized from veins better in these two channels.

Also several classifiers including Kmeans, fuzzy clustering, SVM and LDA are evaluated for classification of arteries and veins. For training and evaluation of the classifiers we use DRIVE database which include 20 train and 20 test images. The training set is divided into two sets of 10 images. The first 10 images are used for training the classifiers. The other 10 images are used for determination of parameters, features and the best classifier. Finally the best classifier is trained with the set of 10 training images and is applied to

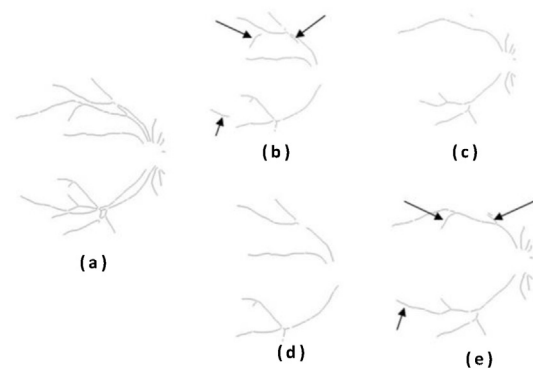


Fig. 8. (a) Major arteries and veins classified by the proposed method. (b) Artery map before post-processing. (c) Vein map before post-processing. (d) Artery map after post-processing. (e) Vein map after post-processing. As can be seen some incorrectly labeled vessels are returned to their true map using connectivity information and structural features of vascular tree.

20 test images. The classifier assigns an artery or vein label to each sub vessel segment.

All vessel points in each initial vessel segment are of the same class. Thus number of pixels of that vessel segment which are artery or vein are counted and dominant label is assigned to every pixel of that vessel segment. The total accuracy of the classification method is measured by:

$$Accuracy = \frac{n_c}{n_c + n_i} \times 100 \quad (10)$$

where n_c is the number of vessel points that receive a correct artery or vein label by the classifier and n_i is the number of incorrectly labeled pixels.

3.3. Post-processing

In the previous step vessels thicker than 3 pixels all over the retinal images have been classified into arteries and veins. The output of the previous step is two binary images of vessel maps for arteries and veins. As can be seen from Fig. 8 there are some vessel segments that are not connected to the detected artery or vein tree. Indeed these vessel segments were incorrectly labeled and must return to the other vessel map. In this section we attempt to correct the label of such vessels based on the labels of their adjacent vessel segments or other vessels that are connected to them.

The post-processing stage includes two steps. First connected vessels of the same type are found using the structural knowledge at cross over and bifurcation points. The structural knowledge of the retinal vascular tree is used in form of two rules. The first rule demonstrates that if three vessel segments are connected to each other through a bifurcation point, then all of the three vessels must be of the same type. The second rule demonstrates that if two vessels cross each other, one is artery and the other one is vein. In the second step, for each detected sub tree of artery or vein, number of vessel points labeled as arteries and veins are counted and the dominant label in that sub tree is found. If the number of vessel pixels with the dominant label exceeds a threshold then the dominant label of that vessel sub tree is assigned to all vessel points of that tree.

To find sub trees of connected vessels, we start from vessel segments labeled as arteries or veins in the previous step. We also need the skeleton of the entire vessel network to discriminate cross over and bifurcation points. Consider vessel segment I as the first previously classified vessel segment for which the respective connected tree is going to be found. We must find all other vessel segments which are connected to this segment I directly or indirectly. To find these vessels, two end points of vessel I are investigated. If the end

Table 1
Explains the best features extracted from vessels.

F#	Feature description for each sub vessel
1	Mean value for pixels of the skeleton of each sub vessel in red image
2	Mean value for pixels of the skeleton of each sub vessel in green image
3	Mean value for all pixels of sub vessel in red image
4	Mean value for all pixels of sub vessel in Green image
5	Variance value for all pixels of sub vessel in A channel of LAB color space
6	Variance value for all pixels of sub vessel in B channel of LAB color space
7	Difference of mean of intensity values on pixels of wall and centerline pixels of vessels in Red channel
8	Difference of mean of intensity values on pixels of wall and centerline pixels of vessels in Luminance channel

point is a bifurcation point as was indicated by rule number one the other two vessel segments connected to I are of the same type as I and belong to the sub tree. However we only add those vessels for which the classifier has assigned a label. If the end point is a cross over, then one of the vessel segments adjacent to I is of the same type as I . To find that vessel segment from the three adjacent vessels, we pick the vessel segment which has the same direction as vessel I . The procedure is then repeated for each vessel segment that has been added to the sub tree until there is no more vessel segment to be added. Then the dominant label for the sub tree of vessels is found. If the number of vessel points of the same type as the dominant label is higher than threshold θ (above 75%) then the dominant label is propagated through the vessel sub tree.

The procedure indeed attempts to find artery or vein sub trees of major vessels with the maximum connected vessels possible using the structural knowledge and revise the vessel labels considering the dominant label in each sub tree. Also note that during the extraction of vessel centerline points, cross over points turn into two bifurcation points very close to each other. Therefore instead of counting the number of adjacent pixels to each centerline points a window is located on the centerline point and if the number of intersections of vessel with that window is four, the point might be a cross over. In this paper we have used a circular window of radius 8 pixels.

As seen in Fig. 8 some of the vessel segments that has been labeled mistakenly and thus are not connected to the detected artery or vein map are returned to their true vessel map.

4. Experimental results

In this section we study the performance of our proposed method on two datasets described previously and attempt to perform a fair comparison with other approaches proposed in the literature for artery/vein classification of retinal blood vessels.

Automated classification of blood vessels has been much investigated for classification of major vessels aimed at AVR measurement. For AVR measurement as was explained previously, major vessels in a circular region around the (O.D.) and of a certain distance from O.D. are extracted and classified into arteries and veins. Major vessels are usually considered as those vessels which directly exit from optical disk. In some methods only main arteries and veins running parallel to each other are taken into account for AVR measurement [16]. Other methods consider major vessels as those which are distinguishable by ophthalmologist using only their color features [12]. For classification of major vessels using supervised methods, classifier is trained using arteries and veins which have been marked by ophthalmologists only considering their color features. Indeed, involving secondary vessels greatly decrease the recognition rate of major vessels since small vessels are hardly recognizable by their color.

Also classification rate of arteries and veins is affected by the region where vessels are located. V'azquez et al. [14] discussed the artery/vein classification rate of vessels in regions of different distances from O.D. to determine the best region for AVR calculation

in their paper. They concluded that the region of radius 2.5 times the optic disk radius is the best one for AVR measurement for their database. In addition to these problems, another factor that might make the comparison of different methods difficult is the recognition rate of the method used for extraction of vessels. In Table 2, the most recent methods for classification of retinal blood vessels are compared.

In this paper, we attempted to classify the vessels thicker than three pixels all over the retinal image with high recognition rate while maintaining the recognition rate in ROI high for AVR calculation. Classification of major vessels all over the images will potentially allow for developing an automated method for classification of the complete vascular network since small vessels are labeled by tracing them back to the major vessels. For vessel segmentation we used the Soares [31] code which is publicly available. The accuracy of their method on DRIVE database is 96%.

For separation of arteries from veins we investigate both unsupervised and supervised classification. Unsupervised classification of vessels is of high interest for ophthalmologist due to independent classification of vessels in each image and preventing the method from being favored toward a particular dataset. However, high variation of illumination through each image makes it hard for clustering based methods to efficiently classify vessels all over the retinal image. Some methods such as [8,14] divide the O.D. based images into four symmetric partitions and perform clustering on each region independently to lessen the effect of non-uniform illumination. However, in these methods it is necessary that each region includes at least a pair of artery and veins. Since our data is not necessarily O.D. centered, instead of dividing the image into four regions based on O.D. we use overlapped circular regions as is shown in Fig. 9 and perform clustering in each region independently. Combination of results in different circular regions improves the classification result comparing to when clustering is performed on the whole image. Finally after comparing the artery/vein classification rate obtained by support vector machine

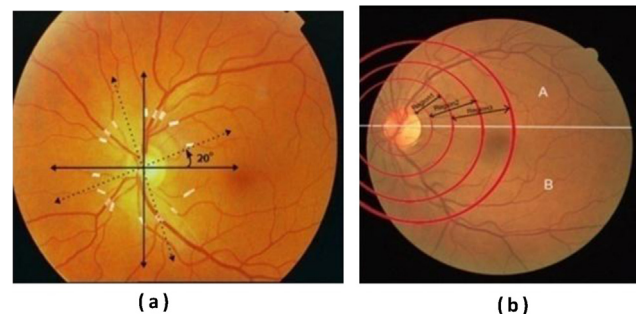


Fig. 9. (a) Retinal image is partitioned into four symmetric regions based on O.D. Each partition is rotated 20° and clustering is applied [14]. (b) Retinal image is partitioned into circular regions around O.D. and also into two upper and lower regions A and B. Clustering is applied on each circular region and also on two consecutive circles in each region A and B independently. Final results are combined based on major voting.

Table 2
 Result of most recent automated methods for retinal blood vessel classification against public available databases.

	Method	Database	Accuracy	Description
Vazquez et al. [14]	K-means clustering	VICAVR	91%	OD based images
Niemeijer et al. [12]	kNN classification + 12 features	DRIVE	0.88 ROC curve area	Major vessels separated by ophthalmologist + manually segmented vessels
Niemeijer et al. [15]	LDA classification + 27 features	65 High resolution O.D. centered images	0.84 ROC curve area	All vessels in ROI of 1–1.5 dist diameter
Muramatsu et al. [16]	LDA classification + 5 features	DRIVE	93%	Vessels thicker than three pixel in ROI of 0.5–1 DD from OD margin + sensitivity for vessel extraction method is 87%

Table 3
 Result of the proposed method for vessels in the entire retina.

Database	Average	Min	Max	Arteries	Veins	Description
DRIVE	84.05	70.71	93.91	82.65	85.74	Vessel thicker than three pixel all over retina + vessel extraction accuracy is 96%
Khatam	80.10	68.55	88.89	71.18	88.13	

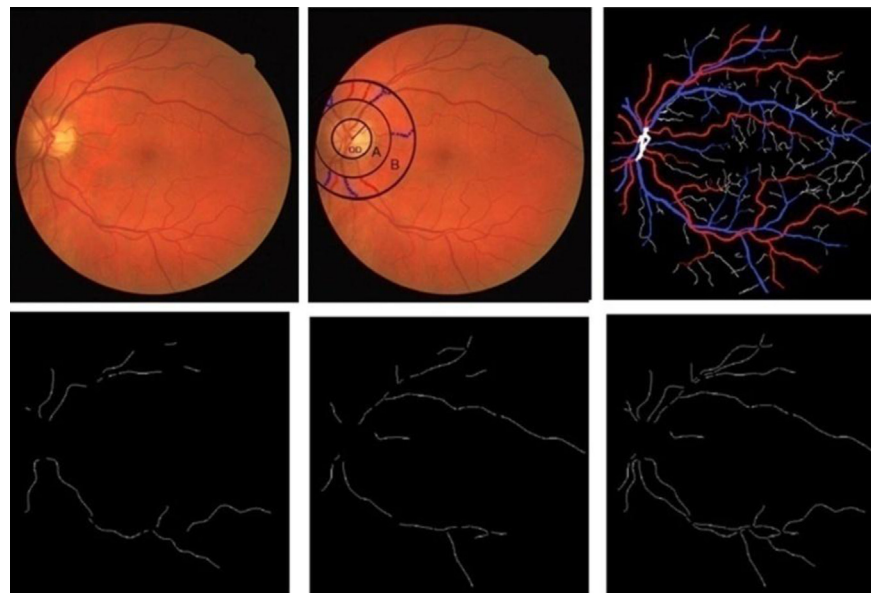


Fig. 10. (a) A sample RGB image. (b) Result of classification of vessels in AVR measurement zone. (c) Manual labels for arteries and veins excluding small vessels. (d) Artery map produced by the proposed method. (e) Vein map produced by the proposed method. (f) All vessels thicker than three pixels classified by the automated proposed method.

classifier, linear discriminant analyzer classifier and fuzzy c-means clustering, LDA is chosen as the best classifier for our method.

For each feature vector, appropriate label is assigned using the reference labels provided by our ophthalmologist and LDA is trained. The 20 test images are then preprocessed similarly to training images and one time trained LDA classifier is applied to them. Since all vessel points in a vessel segment are of the same type, the dominant label for each vessel segment is obtained and assigned to every pixel of that vessel segment. Then the post-processing step is applied to images to revise the incorrectly labeled vessels using labels of their connected vessels which raise the accuracy up to 1.40%. Final results on centerline pixels of major vessel which include approximately 48% of all centerline pixels in our method are shown in Table 3. In Table 3 min/max refers to the minimum/maximum percentage of vessels correctly classified in an image of a dataset; average is the average of correctly classified vessels in the whole dataset; and arteries/veins is the average percentage of correctly classified arteries/veins in the whole dataset. Table 4 shows number of images for which we have rise, fall or no change in accuracy after post processing. Also Fig. 10 shows a

sample retinal image and the respective outputs of the proposed method.

Niemeijer et al. method [15] achieved the 84% roc curve area for classification of all vessels in ROI while Mutamatsu et al. [16] demonstrated 93% accuracy on vessels thicker than three pixels. In comparison, our method, is trained with all vessels thicker than three pixels all over the retina and include many secondary vessels. In addition the images used in this paper are not O.D. based thus they contain less major vessels comparing to O.D. based images. Also we used an automated method for extraction of vessels which

Table 4
 Shows number of images for which we have rise, fall or no change in accuracy after post processing.

Database	Image numbers with			Total images
	Rise of accuracy	Fall in accuracy	No change in accuracy	
DRIVE	8	2	10	20
Khatam	7	1	5	13

Table 5
Result of the proposed method in the region of interest for AVR measurement.

Database	Average	Min	Max	Arteries	Veins	Description
DRIVE	90.16	76.22	100	90.85	89.83	Vessels thicker than three pixels in ROI of
Khatam	88.18	78.97	100	95.21	81.26	0.5–1 DD from O.D. margin

extracted 96% of vessels correctly. Finally we have achieved high recognition rate without increasing number of training samples or features. Indeed our method divides vessel segments and classifies sub vessels segments instead of vessel points. Using mean and variance of color and light over more pixels comparing to one pixel intensity features or profile based features, lead to smaller feature vector with features robust to outliers.

We also evaluate the classification results in regions of different diameters from O.D. to determine the best region for AVR determination. For this purpose optical disk is manually detected by our ophthalmologist, and artery/vein classification rate is calculated for vessels in regions of different radii from O.D. Results demonstrate that the best region has a distance of 0.5 to 1 disk diameter (DD) from O.D. margin. Final results in the region of interest with radii 0.5 to 1 optical disk from O.D. margins are shown in Table 5.

To test the robustness of the proposed method, we evaluate the method on a second database of high resolution images consisting of 13 retinal images. All the images are preprocessed similarly to the images in DRIVE database. Since the DRIVE images are of a particular resolution, applying the trained classifier to the Khatam images might not produce adequate results. Thus the high resolution images are down sampled with the factor of 4. One time trained LDA classifier is applied to the images without changing any parameters previously used for DRIVE database. Finally images are post processed by the method explained in Section 2.3. Although, training the classifier with the images from the same database and adjusting the parameters will produce better results due to same illumination condition while imaging for images in the same database, final results achieved using the LDA classifier trained with DRIVE images still are considerably good.

Finally we believe that, the problem of non-uniform illumination and variation of color between images still may be improved to achieve better results. This may be achieved by taking advantage of local information through partitioning the image. Also exploiting image segmentation methods which maintain high degree of vessel connectivity in vascular network and adding a conflict resolution step may help considerably for developing an efficient automated method for propagating the labels of major vessels toward small vessels. Finally using features that capture the characteristics of centerline points of vessels as well as variation of lightness and color along the vessels more precisely will help to further improve the method.

5. Conclusions

In this paper a novel automated and structural method for classification of retinal blood vessels into arteries and veins has been presented. The proposed method classifies the major vessels in the entire image while maintain high classification rate for vessels in region of interest for AVR measurement.

We have employed several image enhancement methods to deal with the main problems in classification of retinal blood vessels which are non-uniform illumination and color variation through images. In feature extraction step, we attempted to decrease number of the training sample points and features by classification of vessel segments instead of vessel points along with investigating features in RGB, HSL and LAB color spaces.

We also employed connectivity information at cross over and bifurcation points to revise some incorrectly labeled vessels.

Evaluation of our method against two different datasets of retinal images demonstrates the good performance and robustness of our method. While we trained the classifier with vessels in the entire image we could achieve high classification rate in region of interest for AVR measurement. Our method may be used for AVR measurement and as a start map for extending the artery/vein classification to the entire vascular network.

We believe further investigation of local information and vessel connectivity, using larger datasets of images and better image enhancement techniques for color normalization inter images may help for improving the method.

Acknowledgments

The retinal images of Khatam dataset have been provided in Khatam-Al-Anbia eye hospital of mashhad, Iran. Data used in this paper has been annotated manually under supervision of Dr. Touka Banaee at Mashhad university of Medical science, Iran.

References

- [1] Frame AJ, Undrill PE, Cree MJ, Olson J, McHardy KC, Sharp PF, et al. A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms. *Comput Biol Med* 1998;28:225–38.
- [2] Nguyen TT, Wang JJ, Wong TY. Retinal vascular changes in pre-diabetes and pre-hypertension, new findings and their research and clinical implications. *Diabetes Care* 2007;30:2708–15.
- [3] Ikram MK, de Jong FJ, Vingerling JR, Witteman JCM, Hofman A, Breteler MMB, et al. Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The Rotterdam study. *Invest Ophthalmol Vis Sci* 2004;45(July (7)):2129–34.
- [4] Sun C, Wang JJ, Mackey DA, Wong TY. Retinal vascular caliber: systemic, environmental, and genetic associations. *Survey Ophthalmol* 2009;54(1):74–95.
- [5] Yousif AAA, Ghalwash AZ, Hongyu A. A comparative evaluation of preprocessing methods for automated detection of retinal anatomy. In: *Proceedings of the fifth international conference on informatics & systems (INFOS'07)*. 2007.
- [6] Foracchia M, Grisan E, Ruggieri A. Luminosity and contrast normalization in retinal images. *Med Image Anal* 2005;9(June (3)):179–90.
- [7] Grisan E, Ruggieri A. A divide et impera strategy for automated classification of retinal vessels into arteries and veins. In: *25th international conference of the IEEE engineering in medicine and biology society*, vol. 1. 2003. p. 890–3.
- [8] Li H, H. W, Mong Lee L, Ghoneim AS. A piecewise Gaussian model for profiling and differentiating retinal vessels. In: *ICIP03*, vol. 1. 2003. p. 1069–72.
- [9] Jelinek HF, Dpardieu C, Lucas C, Cornforth DJ, Huang W, Cree MJ. Towards vessel characterization in the vicinity of the optic disc in digital retinal images. In: *Proceedings of the image and vision computing New Zealand*. 2005.
- [10] Kondermann C, Kondermann D, Yan M. Blood vessel classification into arteries and veins in retinal images. In: *Med Imag 2007: Image Process*, vol. 6512. 2007. p. 651247–9.
- [11] Rothaus K, Jiang X, Rhiem P. Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. *Image Vis Comput* 2007;27(7):864–75.
- [12] Niemeijer M, van Ginneken B, Abramoff MD. Automated classification of retinal vessels into arteries and veins. In: *Karssemeyer Nico, Giger Maryellen L, editors. Medical imaging 2009: computer-aided diagnosis*. Proceedings of the SPIE, vol. 7260. 2009. p. 72601F–172601F.
- [13] Vazquez SG, Barreira N, Penedo MG, Penas M, Pose-Reino A. Automated classification of retinal vessels into arteries and veins. In: *7th international conference biomedical engineering (BioMED 2010)*, vol. 1. 2010. p. 230–6.
- [14] Vazquez SG, Barreira N, Penedo MG, Saez M, Pose-Reino A. Using retinex image enhancement to improve the artery/vein classification in retinal images. In: *ICIAR 2010, Part II, LNCS*, vol. 6112. 2010. p. 50–9.
- [15] Niemeijer M, Xu X, Dumitrescu AV, Gupta P, Ginneken BV, Folk JC, et al. Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. *IEEE Trans Med Imag* 2011;30(November (11)):1941.
- [16] Muramatsu C, Hatanaka Y, Iwase T, Hara T, Fujita H. Automated selection of major arteries and veins for measurement of arteriolar-to-venular diameter ratio on retinal fundus images. *Comput Med Imag Graph* 2011;35:472–80.
- [17] Tramontan L, Grisan E, Ruggieri A. An improved system for the automated estimation of the arteriolar-to-venular diameter ratio (AVR) in retinal images. In:

- Proceedings of the 30th annual international conference of the IEEE engineering in medicine and biology society. 2008. p. 3550–3.
- [18] Saez M, Gonz'alez-V'azquez S, Gonz'alez-Penedo M, Ant'on'ia-Barcel'o M, Pena-Seijo M, Coll de Tuero G, et al. Development of an automated system to classify retinal vessels into arteries and veins. *Comput Methods Programs Biomed* 2012. <http://dx.doi.org/10.1016/j.cmpb.2012.02.008>.
- [19] Hubbard LD, Brothers RJ, King WN, Clegg LX, Klein R, Cooper LS, et al. Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology* 1999;106(12):2269–80.
- [20] Knudtson MD, Lee KE, Hubbard LD, Wong TY, Klein R, Klein BEK. Revised formulas for summarizing retinal vessel diameters. *Curr Eye Res* 2003;27(3):143–9.
- [21] Aguilar W, Martinez-Perez M, Frauel Y, Escolano F, Lozano M, Espinosa-Romero A. Graph-based methods for retinal mosaicing and vascular characterization. *Curr Eye Res* 2003;27(3):143–9.
- [22] Wu PCh, Cheng FCh, Chen Y-K. A weighting mean-separated sub-histogram equalization for contrast enhancement. In: *International conference on biomedical engineering and computer science (ICBECS)*. 2010.
- [23] Tsai ChM, Yeh ZM. Contrast enhancement by automated and parameter-free piece-wise linear transformation for color image. *IEEE Trans Comput Electr* 2008.
- [24] Mirsharif G, Tajeripour F. Investigating image enhancement methods for better classification of retinal blood vessels into arteries and veins. In: *Artificial Intelligence and Signal Processing 2012*. 2012.
- [25] Jobson DJ, Rahman Z, Woodell GA. Properties and performance of a center surround retinex. *IEEE Trans Image Process* 1997;6(3).
- [26] Jobson DJ, Rahman Z, Woodell GA. A multi-scale retinex for bridging the gap between color images and human observation of scenes. *IEEE Trans Image Process* 1997;6(3).
- [27] Vlachos M, Dermatas E. Multi-scale retinal vessel segmentation using line tracking. *Comput Med Imag Graph* 2010;34(April (3)):213–27.
- [28] Zhang B, Zhang L, Zhang L, Karray F. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Comput Biol Med* 2010;438–45.
- [29] Shu B, Larn Y, Yong H. A novel vessel segmentation algorithm for pathological retina images based on the divergence of vector fields. *IEEE Trans Med Imag* 2008;27(February (2)):237.
- [30] Poureza R, Poureza H, Banae T. Segmentation of blood vessels in fundus color images by radon transform and morphological reconstruction. In: *Third international workshop on advanced computational intelligence*, August 25–27. 2010.
- [31] Soares JV, Leandro JJ, Cesar J'unior RM, Jelinek HF, Cree MJ. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE Trans Med Imag* 2006;25(September (9)):1214–22.
- [32] Ahmed M, Ward R. A rotation invariant rule-based thinning algorithm for character recognition. *IEEE Trans Pattern Anal Mach Intell* 2002;24(12):1672–8.
- [33] Rockett P. An improved rotation-invariant thinning algorithm. *IEEE Trans Pattern Anal Mach Intell* 2005;27(10):1671–4.
- [34] Mirsharif G, Tajeripour F, Sobhanmanesh F, Poureza H, Banae T. Developing an automated method for separation of retinal blood vessels into arteries and veins. *Int Conf Comput Knowl Eng* 2011.
- [35] Staal JJ, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge based vessel segmentation in color images of the retina. *IEEE Trans Med Imag* 2004;23:501–9.

Qazaleh Mirsharif has received BS from computer engineering department at azad university of Mashhad in 2008 and a M.Sc. in artificial intelligence from computer engineering department at Shiraz university of Iran in September 2012. Her research interest includes image and video processing.

Farshad Tajeripour received the B.Sc. and M.Sc. degrees in electrical engineering from Shiraz University, in 1994 and 1997, respectively, and the Ph.D. degree in electrical engineering from the Tarbiat Modarres University of Tehran, in 2007. In 2007, he joined the Department of Electrical and Computer Engineering at Shiraz University, Shiraz, as an Assistant Professor. His research interests include digital image processing, machine vision, medical image processing, signal processing, and the design and implementation of vision based inspection systems.

Hamidreza Poureza received the B.Sc. in electrical engineering from Ferdowsi university of Mashhad in 1989, the M.Sc. degree in electrical engineering from Amirkabir university of Technology (Tehran Polytechnic) in 1993 and a Ph.D. degree in computer engineering from Amirkabir university of Technology (Tehran Polytechnic) in 2003. He is currently an associate professor in computer engineering department at Ferdowsi university of Mashhad, Iran. His research interests include Computer vision, signal processing, hardware design and intelligent transportation systems.