Journal of Innovative Optical Health Sciences
Vol. 13, No. 1 (2020) 1950021 (34 pages)

DOI: 10.1142/S1793545819500214

Classifiers fusion for improved vessel recognition with application in quantification of generalized arteriolar narrowing

Xiaoxia Yin*,‡, Samra Irshad† and Yanchun Zhang†
*Cyberspace Institute of Advanced Technology
Guangzhou University
Guangzhou 510006, P. R. China
†Institute for Sustainable Industries and Liveable Cities
Victoria University
Melbourne, Australia
‡xiaoxia.yin@gzhu.edu.cn

Received 14 April 2019
Accepted 6 August 2019
Published 25 November 2019

This paper attempts to estimate diagnostically relevant measure, i.e., Arteriovenous Ratio with an improved retinal vessel classification using feature ranking strategies and multiple classifiers decision-combination scheme. The features exploited for retinal vessel characterization are based on statistical measures of histogram, different filter responses of images and local gradient information. The feature selection process is based on two feature ranking approaches (Pearson Correlation Coefficient technique and Relief-F method) to rank the features followed by use of maximum classification accuracy of three supervised classifiers (k-Nearest Neighbor, Support Vector Machine and Naïve Bayes) as a threshold for feature subset selection. Retinal vessels are labeled using the selected feature subset and proposed hybrid classification scheme, i.e., decision fusion of multiple classifiers. The comparative analysis shows an increase in vessel classification accuracy as well as Arteriovenous Ratio calculation performance. The system is tested on three databases, a local dataset of 44 images and two publically available databases, INSPIRE-AVR containing 40 images and VICAVR containing 58 images. The local database also contains images with pathologically diseased structures. The performance of the proposed system is assessed by comparing the experimental results with the gold standard estimations as well as with the results of previous methodologies. Overall, an accuracy of 90.45%, 93.90% and 87.82% is achieved in retinal blood vessel separation with 0.0565, 0.0650 and 0.0849 mean error in Arteriovenous Ratio calculation for Local, INSPIRE-AVR and VICAVR dataset, respectively.

Keywords: Hypertensive retinopathy; retinal vessel classification; optic disk; arteriovenous ratio; region of analysis; support vector machine.

1. Introduction

Cardiovascular diseases including coronary heart disease, stroke and hypertension are characterized by the deviations in blood vascular structure.[1,2] In hypertension, arteries are altered from the regular pattern and the inner lining of arteries is damaged and as a result, they become thick and stiff.[3] Due to this thickness of artery-walls in hypertension, the normal blood flow pressure is affected.[4] Along with other body organs including heart and kidney, the presence of hypertension also affects eye and leads to several ocular disorders including Hypertensive Retinopathy (HR).[5] In HR, both the vascular and nonvascular structures in retina are deteriorated. However, the alteration of retinal vessel width is considered as an earliest biomarker of HR.[6] Particularly, the width of arteries is narrowed in initial stages, that is why "arteriolar narrowing" is placed at initial stage in all the three scales proposed so far for HR diagnosis.[7–10] For assessment of arteriolar narrowing, a parameter called Arteriovenous Ratio (AVR), suggested by Stokoe and Turner,[11] is used. It is the ratio of average diameter of retinal Arterioles (arteries) to average Venules (veins) diameter and its calculation involves the use of two other parameters known as Central Retinal Artery Equivalent and Central Retinal Venular Equivalent.[12,13] The deviation of this parameter from a normal range indicates the presence of HR so this biomarker is considered crucial for HR severity assessment. In order to quantify the AVR, the ophthalmologists examine the internal structure of retina using the photographs obtained via different imaging modalities, i.e., Ophthalmoscopy, Fluorescein Angiography and Fundus Photography, however, fundus imaging is the only non-invasive imaging method that provides a mode for extensive visualization of blood vessels and other structures in retina. Moreover, according to the studies,[14–16] the fundus camera has the additional capability of improving retina visualization and consequently the detection of retinal diseases. Figure 1 illustrates a sample retinal fundus image taken from local database with several anomalies that occur in HR.

Generally, AVR is manually estimated by ophthalmologists through visual screening of retina in fundus image which is a time-taking and labor-intensive process. The ophthalmologist first manually performs vessel classification in retinal photographs and then estimates AVR. To assist the ophthalmologist to speed up the detection process, the digital image processing techniques and computer vision tools are being used for efficient and timely diagnosis of HR.[17–22]

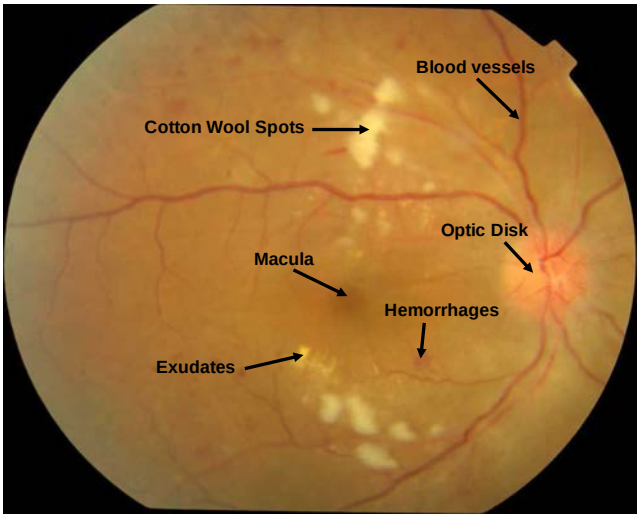


Fig. 1. Sample retinal image with structures and abnormalities marked with arrow.

Automated HR detection systems generally consist of three modules, i.e., retinal vessel, (a) segmentation, (b) classification and (d) width estimation. The accurate segmentation and classification of retinal vessels directly affects the retinal width estimation which is later used for AVR calculation. A vast number of computerized retinal analysis studies have focused on vessel segmentation.[23–25] There are a few researches that have focused on detection of vessel bifurcations and crossover points.[26,27] As far as retinal vessel classification is concerned, there are existing methods that recognize vessels as veins and arteries either automatically or semi-automatically.[28,29] The most prominent visual difference between retinal veins and arteries is their color. Arteries are light in color due to the abundance of oxygenated hemoglobin and this fact has led to the use of intensity-based features for vessel recognition. The pioneer study was proposed by Grisan and Ruggeri[88] for arteriovenous classification in which retinal image was divided into four quadrants based on optic disk center and vessel segments found in each quadrant are classified using color features. This quadrant division-based approach is also adopted in later studies.[30] In another method proposed by the same group,[31] several circular regions of different radii around optic disk are assessed to obtain the classes

of vessels in each zone using vessel profile features. A similar approach is proposed by Mirsharif *et al.*[32] with a little variation by dividing the image in four quadrants with further partitions in upper and lower regions.

Some of the methods presented previously also employed inherited vessel-structure sequence information along with color features for retinal vessel classification. A method proposed by Niemeijer *et al.*[33] for classification of retinal vessels incorporated color features and prior structural knowledge of vessels. Another important inherited characteristic of artery-vein is that the same types of vessel never cross each other. This retinal vessel property has also been employed to classify vessels.[34] Furthermore, in the literature, some methods classify complete vessel network while others consider only small vessel segments within specific circular zone for classification.[35,36] However, it is not necessary to classify complete vessel network for calculation of AVR.[3] Graph-based methods have been investigated in Ref. 37. In these methods, graphs are generated using vessel center-lines with vessel junction points and the resultant graphs are classified into arteries and veins. Most of these methods classified the complete vessel network.[38] Same group Dashtbozorg *et al.*[39]; Mendonça *et al.*[40] also proposed methods for AVR calculation. Among the recent methods, Xu *et al.*[41] proposed first- and second-order textural features for differentiation between arteries and veins in retina and Yan *et al.* (2017) used context-dependent features for retinal vessel classification. Quite recently, Relan *et al.* (2017) presented approach for arteriovenous classification that includes the use of a pre-processing module called multiscale self-quotient image method for illumination and lightning correction in retinal images. Currently, the deep learning network has also been widely and successfully applied in the retinal image processing and segmentation.[42–44]

Although, the advancement in technology has paved the path for robust and reliable development of computer algorithms for HR detection through assessment of retinal vessels, however, it is observed that the methods previously proposed are tested on healthy retinal images, that do not contain pathological structures. The disease progression makes the retinal vessels prone to many different kinds of pathologies, e.g., vessel tortuosity, hard and soft exudates, branch retinal artery and vein occlusion, sheathing of vessels, focal arteriolar narrowing and optic disk swelling, which deteriorate the width and intensity of vessels. Moreover, HR progression has different effect on vein and artery; comparatively arteries are observed to be more affected in presence of HR. Figure 2 shows visualization of such cases in enlarged slices of six images, taken from retinal fundus database acquired from Armed Forces Institute of Ophthalmology, Pakistan. Development of new "ghost" vessels and vessel fading is observed due to occlusion in Fig. 2(a). As shown in Fig. 2(b), retinal arteries are hardly visible (e.g., the ones indicated by circle arrow) due to occlusion and sheathing of vessels with appearance of cotton wool spots in fundus area. Optic disk swelling can be noted from Figs. 2(b) to 2(d). Branch retinal vein occlusion is shown in Fig. 2(c) which makes the vessel appearance deteriorated. It is observed that these pathologies pose a serious difficulty in retinal vessel classification phase, since vessel intensity is the major character that is used to classify vessels. The performance of retinal vessel classification methods on pathologically diseased images needs to be evaluated so that a robust solution can be devised which is invariant to presence of pathological changes. This paper attempts to provide a system for reliable HR diagnosis using AVR parameter that is robust to presence of pathological structures in fundus images. Furthermore, the features extracted for retinal vessel classification may contain redundant information. For this purpose, we exploited two feature ranking strategies for feature selection. Feature selection is a technique to obtain dimensionally reduced feature set by retaining only those features which are truly relevant for predicting the outcome. There are a number of ways for selection of most significant of most features from a large feature set. Most widely used among these ways are; (a) Wrapper methods: Feature selection based on association between predictors (features) and responses (target labels) before applying machine learning algorithm and (b) Filter methods: Feature selection after applying machine learning techniques. In the first type, significant features are filtered out from the original dataset by evaluating the relevance of individual attributes (features) with the target classes. The criterion according to which the feature-target relationship is measured depends upon the specific feature filtering algorithm. Whereas in second type of feature selection techniques, attribute selection is conducted on the

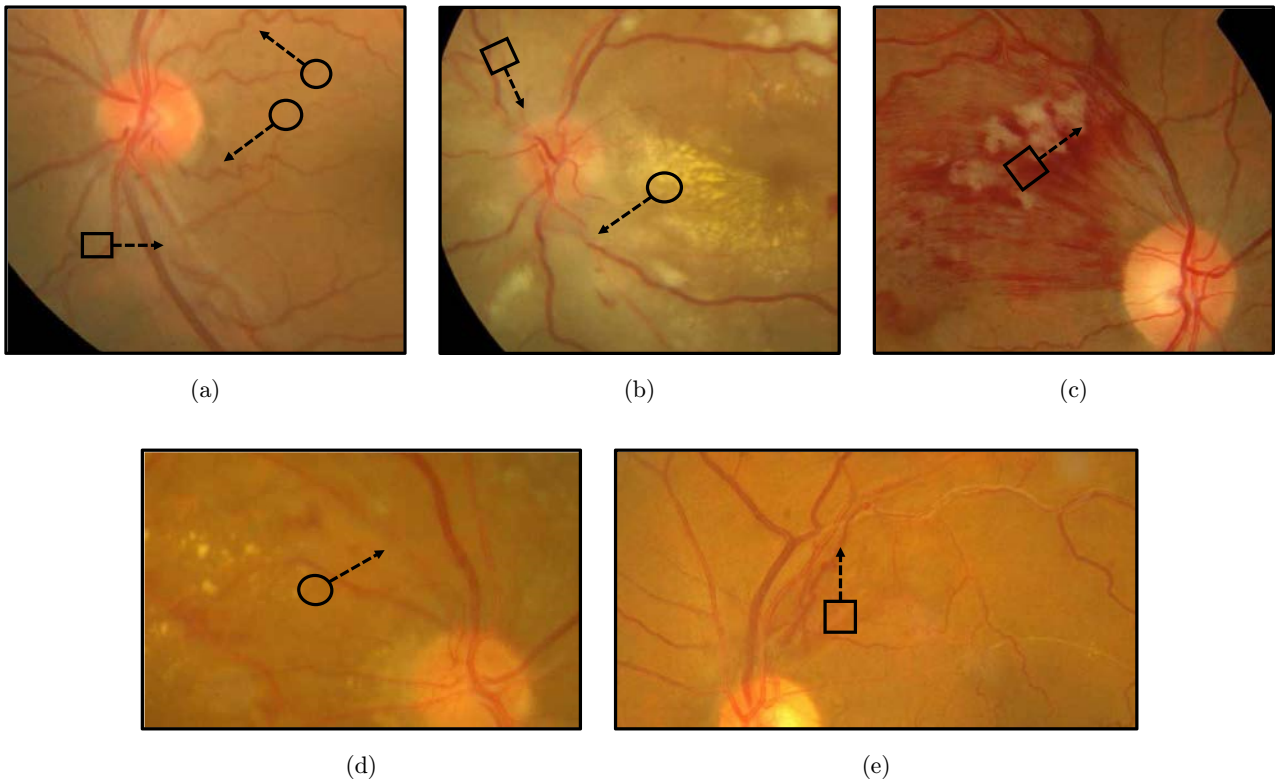


Fig. 2. Image slices showing pathologies (circle and square arrows indicate different abnormalities). (a) Arrow with circle: vessel tortuosity and Neovascularization, Arrow with rectangle: Retinal artery occlusion, (b) Arrow with circle: Shrinkage and occlusion of retinal artery, Arrow with square: Optic Disk Swelling, (c) Arrow with square: Branch Retinal Vein Occlusion, (d) Arrow with circle: Branch Retinal Artery Occlusion and (e) Arrow with square: Sclerotic retinal arteries.

basis of performance of feature in predicting the target class after it has been fed to classifier. Using this technique for feature selection, a feature may show optimal performance for a certain classifier but fails to perform optimally for other classifiers. Effect of filter-based feature selection methods for anomaly detection is examined in depth in many medical applications[45–47] and promising results are reported. The impact of feature ranking methods in relation with decision mapping of machine learning algorithms for Artery/Vein (A/V) classification needs to be evaluated.

In addition to the feature selection process, the class prediction performance of retinal vessels also depends significantly on the particular machine learning algorithm. In methods proposed by, Relan *et al.*[48–50] various predictive models are used for retinal vessel classification: Gaussian Mixture Model, Expectation-Maximization (GMM-EM) classifier, Least Square-Support Vector Machine (LS-SVM) and Squared-loss Mutual Information clustering (SMIC). Similarly, Agurto *et al.*,[51]; Vijayakumar *et al.*[52] used SVM classifier and linear regression, respectively, for retinal vessel classification. In all these approaches, single classifiers are used for vessel recognition. A decision-ensemble based on decision trees with bootstrap aggregation is used by Fraz *et al.*[53] for retinal vessel classification, but this ensemble system is created by employing one type of base classifier. In bootstrap method, different predictive models are generated using different subsets of data and then the decisions of all those models is averaged out. Fusion of different classifiers decisions is known to have enhanced performance as compared to single classifier and has been used in many machine learning applications.[54–57] This motivated us to examine the performance of different classifiers combination for retinal vessel classification.

With respect to the issues highlighted above, this paper aims to develop an automatic framework for detection of HR with an improved retinal vessel recognition module which is based on the use of different feature ranking strategies and classification accuracy of various prediction models (k-Nearest Neighbor, Support Vector Machine and Naïve Bayes)

to select optimal feature subset followed by vessel labeling through the proposed hybrid classification approach. Moreover, the proposed system detects HR robustly not only in healthy images, but also in images with multiple pathological changes. This research allows to investigate the impact of proposed feature selection process and decision-fusion of multiple classifiers on labeling of retinal vessels.

The effect of proposed method on retinal vessel differentiation is evaluated using three performance metrics, i.e., Classification Accuracy, Sensitivity and Specificity and the method is tested on three databases, Local Database, INSPIRE-AVR database[33] and VICAVR database.[31] The performance of AVR calculation module is assessed by calculating the mean error between the estimations provided by ophthalmologist and the ones shown by our method. The promising results presented by proposed system shows its capability in lowering the prevalence of HR and can be proved a valuable tool for retinal screening.

Main contributions of this work are:

(a) The robustness of proposed retinal vessel classification method and subsequent HR diagnosis is examined in images with pathological structures. Currently, there are seldom papers that report hybrid classification approaches on retinal fundus photography with HR. The proposed system detects HR robustly not only in healthy images but also in images with multiple pathological changes.
(b) A novel feature selection strategy that includes filter-based methods with the use of three classifiers for selection of optimal feature subset is employed. The feature ranking is carried out using two feature ranking algorithms, Pearson Correlation Coefficient[58] and Relief-F method,[59] to rank the features (extracted for A/V differentiation) and then, employing classification accuracy of three supervised classifiers as a stopping criteria (threshold) to select the optimal feature subset from ranked feature list.
(c) The proposed "labels-combination" framework has been investigated for recognition of retinal vessels. Majority voting technique is used to combine the decision labels obtained from three classifiers, i.e., k-Nearest Neighbor (k-NN), Support Vector Machine (SVM) and Naïve Bayes, for vessel labeling.

This paper contains five sections; introduction and review of previous work is already explained in this section, Sec. 2 explains the methodology adopted for retinal image preprocessing, vascular network extraction, Optic Disk localization and boundary segmentation, determination of region of analysis, feature extraction for vessel recognition and finally blood vessel width calculation for AVR computation. Section 3 focusses on the details of feature selection process and retinal vessel classification followed by experimental results in Sec. 4. Section 5 summarizes the contributions and limitations of this research.

2. Methodology

The flow chart for the proposed methodology is shown in Fig. 3. The retinal image is first acquired via fundus camera and then preprocessed. After that, the retinal vascular network is detected using Gabor filter bank and a binary vessel map is generated. Details of preprocessing and vessel segmentation is given in Secs. 2.1 and 2.2, respectively. Next, the position of optic disk is determined using Laplacian of Gaussian filter with highest vessel density feature. Based on the optic disk boundary, a circular region of interest is defined around optic disk and the vessels within this region are obtained, as described in Secs. 2.3 and 2.4 Vessel junctions i.e., bifurcations and cross-overs, present in the extracted vessel segments are detected and then, differentiated using local variance based method in order to remove cross-overs. In the next step, a set of 81 features is extracted from retinal vessels as explained in Secs. 2.5. The acquired features are then subject to two feature-ranking methods, i.e., Pearson Correlation Coefficient and Relief-F method. Different combinations of ranked features are then fed to three supervised classifiers (k-NN, SVM and Naïve Bayes) and based on the classification accuracy of each classifier; optimal feature subsets are selected and fused together using union operation. Afterwards, these fused feature subsets are given as an input to decision-fusion framework and final labels of retinal vessel segments are obtained. Classified vessel segments are then measured using 2-Dimentional (2D) Euclidean distance transform, as explained in Sec. 2.6. After measuring the widths of vessel segments, AVR is calculated. Details of feature ranking algorithms, classifiers and feature

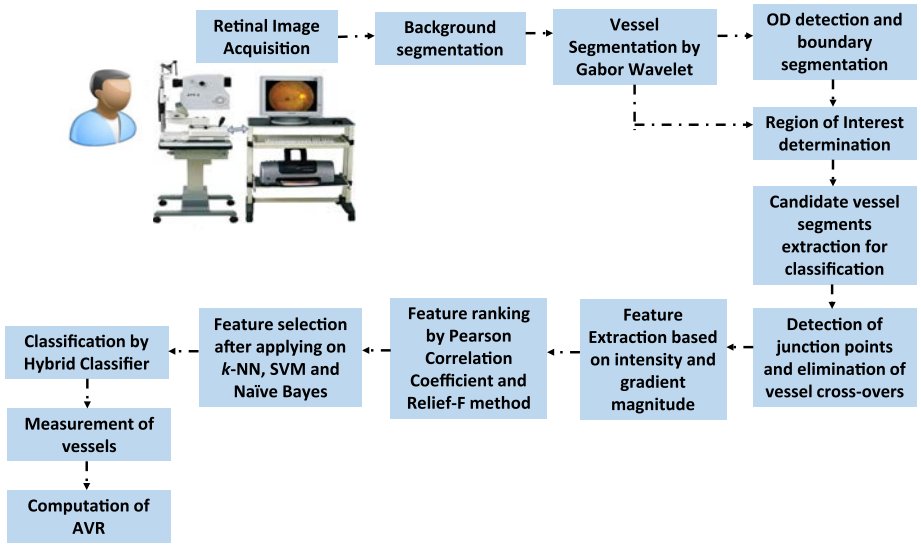


Fig. 3. Flow chart of proposed methodology.

selection method are given in Secs. 3.1 and 3.2, respectively. Experimental results and performance comparison are presented in Sec. 4 followed by discussion in Sec. 5. The proposed methodology involves technically diverse techniques for each step from image preprocessing to calculation of AVR and some steps involve tuning of relevant parameters which affects the final performance of the system. In addition to that, all three image datasets used in this research have different image specifications (size and quality); therefore, in some steps, optimal parameter values may vary. The tuned parameter values are mentioned explicitly where they are different for the three datasets.

2.1. *Preprocessing*

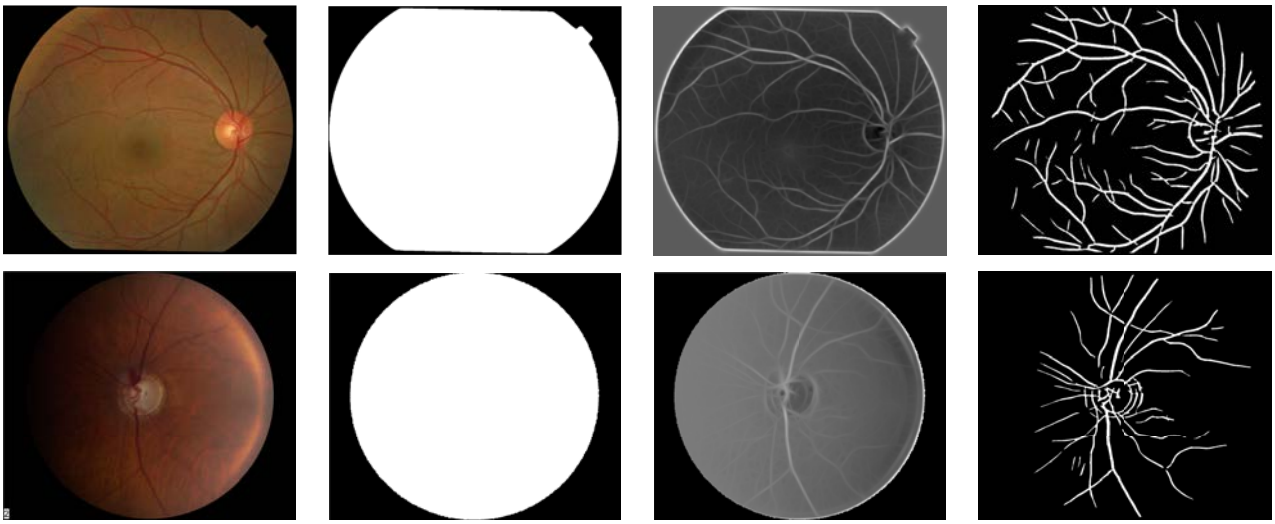
As a preprocessing step, dark background is segmented from digital fundus image in order to reduce the computational complexity. The background is estimated using local mean and variance-based method[60] and binary segmentation mask is formed by thresholding operation, which is dependent on mean of green component of image. The generated background mask is shown in Fig. 4(b) with original image in Fig 4(a), taken from local and INSPIRE-AVR dataset (in first and second row, respectively). Local and INSPIRE-AVR datasets are tuned at a threshold value of 10 whereas for VICAVR dataset, this threshold is set at 30 after background estimation. Regarding threshold selection, we adopt trying and testing techniques, in order to find the binary images with whole retinal information.

2.2. *Vessel extraction*

2D Gabor filter bank is employed here for extraction of vessels.[61] The purpose of using Gabor Wavelet bank is its localization characteristic, due to which the response on small as well large width vessels are captured with greater accuracy. Gabor Wavelet is applied on green channel of image due to the effective discrimination between retinal vessels and fundus area present in this color component. After enhancement of retinal vessel tree, it is thresholded. The Gabor Wavelet is computed for angle spanning from 0° up to 179° at steps of 10° and then the maximum response (MR) is taken. A scale value of 7, 9 and 11 is found to be optimal for local, INSPIRE-AVR and VICAVR dataset, respectively. By varying the angles and scales, Gabor filter bank is formed which enhances the objects in target image according to the set parameters. Figures 4(c) and 4(d) show the enhanced vascular patterns in retinal images using Gabor Wavelet and the binarzied vessel trees, respectively.

2.3. *Identification of the position of Optic Disk (OD)*

OD is a bright circular region in retina from which all the blood vessels emerge. Laplacian of Gaussian (LoG) filter and the highest vessel density property



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

Fig. 4. Vessel preprocessing and extraction. (a) Original retinal image, (b) Background segmented mask, (c) Enhanced retinal vessels using Gabor Wavelet and (d) Binary segmented vessels.

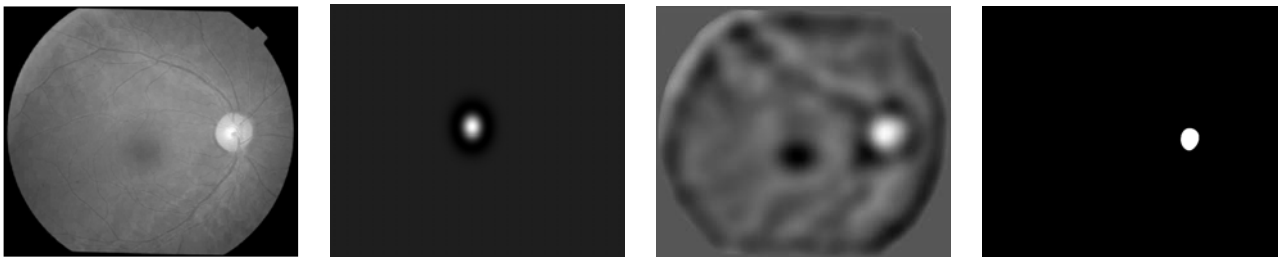
of OD is used in this research to detect the location of OD.[62,63] LoG filter is applied on red channel of RGB image to enhance the location of OD. This template is particularly used because of circular structure of OD and red channel is selected because of clear and discriminating visualization of OD in this channel. After the candidate circular regions have been enhanced using the LoG filter, they are binarized. The threshold value used for binarization is given in Eq. (1).[61]

$$T = 0.6 * m\text{LoG}. \tag{1}$$

This threshold selects pixels having top 60% response from the LoG filtered image and mLoG indicates the maximum value in Gaussian Kernel Processed Image.[62] Preliminary experiments guided the selection of this threshold since it is optimal for all datasets. Red plane of original retinal image, LoG filter, enhanced OD region and binarized OD is shown in Figs. 5(a)–5(d), respectively. Some images in the local dataset contain pathologies like hard exudates and cotton wool spots that have similarity in structural and color properties with OD, so the LoG filtered and subsequent thresholded image may contain more than one OD region. To overcome this issue, vessel density property is incorporated to separate out the OD region from the other segmented portions. After localization of OD position, as illustrated in Fig. 6, its center is determined and boundary is estimated using intensity gradient based technique which we recently proposed in Ref. 63.

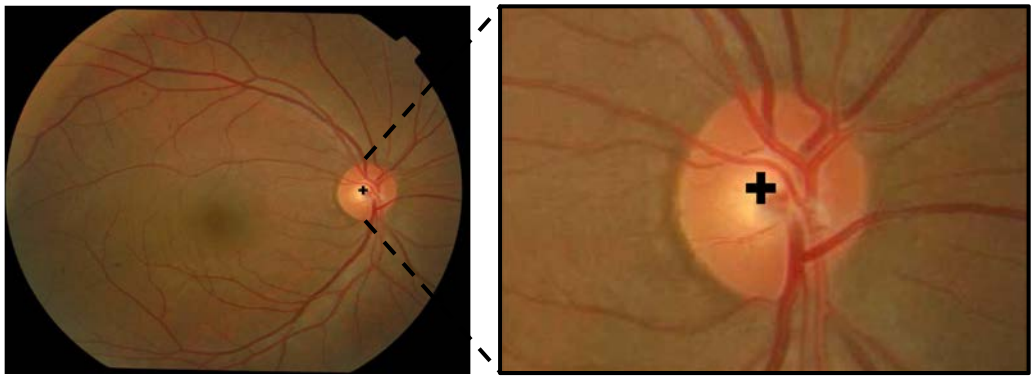
2.4. *Determination of "Analysis Zone" and extraction of vessel segments from vascular tree*

Once the position and boundary of OD is determined, a Region of Analysis (RoI) is identified for



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

Fig. 5. OD detection. (a) Red channel of OD, (b) LoG filter, (c) Enhanced OD region and (d) Binary OD region.

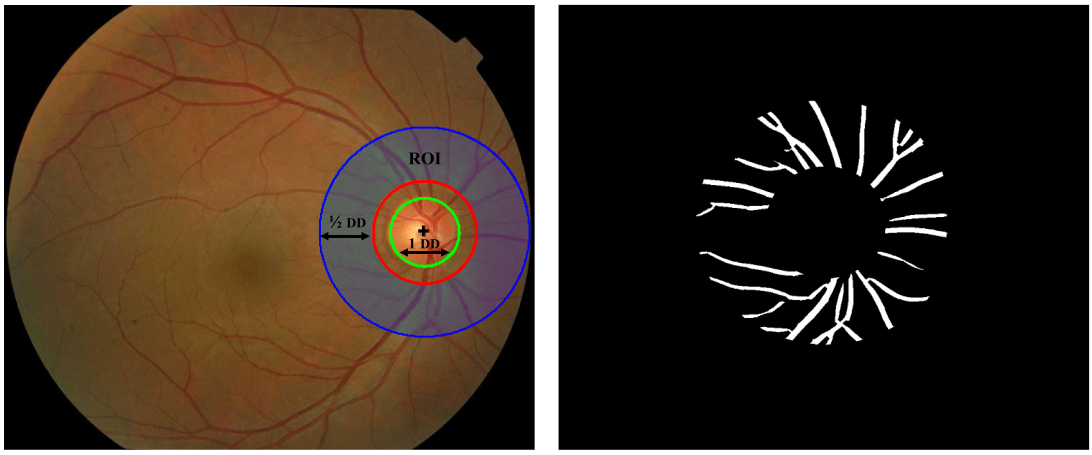


(a) (b)

Fig. 6. OD Localization (a), Position of OD marked with its enlarged version in (b).

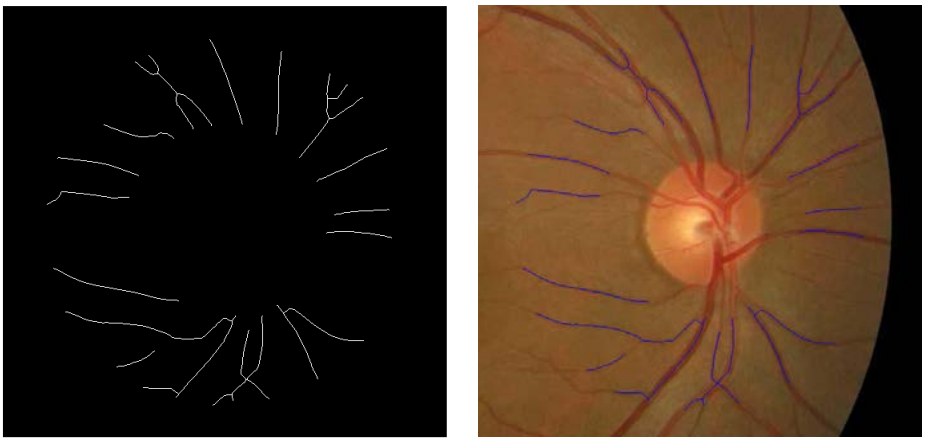
extraction of candidate retinal blood vessels which are to be categorized into artery or vein class. As suggested by Parr and Spears,[64] Knudtson *et al.* (1994) a circular zone that is at a specific distance from OD, is marked and the vessels within this zone are considered for classification and AVR computation purpose. Although, some of the researchers have classified the complete retinal vessel network[37,39,52] but since our goal here is to calculate AVR which can be efficiently calculated by assessing the blood vessels in a specific circular zone around OD, therefore, only the vessel portions within a circular zone are selected. A fixed circular RoI around OD is identified by placing two concentric circles; one at 1/4 Disk Diameter (DD) and another at 1 DD, from OD boundary. Another reason for selection of this zone is to ignore the vessel portions near OD because glial tissue or perivascular sheathing may influence the vessel segments in OD proximity.[11] Figure 7(a) shows the OD boundary and RoI between the two concentric circles and Fig. 7(b) shows the vessels extracted from measurement zone.

After the vessels within RoI are extracted, the next phase is to split those connected vessels into isolated vessel segments by determining the vessel junction points (bifurcations and cross-overs). These landmarks make the vessel classification and measurement task ambiguous. The slices in Fig. 9 illustrates this phenomenon where by examining it can be noticed that the arteriovenous crossing in vessel center-lines appears as one vessel segment, shown in Fig. 9(c). In order to rectify the false landmark, it is important to differentiate between bifurcations and cross-overs. Moreover, even for AVR calculation, the arteries and veins needed to be properly



(a) (b)

Fig. 7. OD boundary. (a) OD boundary marked in green color and RoI is shown between concentric red and blue circle and (b) Retinal vessels segments inside RoI.



(a) (b)

Fig. 8. Skeletonization of retinal binary vessels. (a) Skeletonized binary vessel map and (b) Skeletonized binary vessel segments embedded on original image.

(a) (b) (c)

Fig. 9. Arteriovenous Crossing phenomenon. (a) Slice from original image showing Arteriovenous crossing, (b) Corresponding binary vessel segment and (c) Skeletonized vessel segment.

distinguished, therefore the vessels with crossing points must be isolated into individual vessel segments. In this paper, we adopt a local variance-based method for differentiation between two types of junction points. For determination of these junction points, following steps are implemented.

- First, the binary vessel map is skeletonized and then potential junction points are extracted. Skeletonization is an operation that removes the pixels from edges of objects without destroying the connectivity in an eight-connected scheme and as a result, one-pixel wide center-line vessel structure is extracted.[23] For skeletonization, a level-set algorithm proposed in Rumpf and Telea,[66] is used. The reason for choosing this method is that it gives more smoother and centered structures than other thinning methods. Moreover, center-line vessels avoid pruning branches via this algorithm. Figure 8(a) show the center-line vessels and (b) center-line vessel structure embedded on RGB retinal image.
- Then, this skeleton-vessel image is convolved with a kernel of 3×3 shown in Fig. 10(a) and for each pixel, the number of neighboring pixels are counted. The location of pixels which have three or more than three neighbors is taken. These locations indicate vessel junction points, i.e., bifurcations, arteriovenous crossings or crossing between vessel and capillary. Figure 10(b) shows the pixel localized that have three or more than three neighbors whereas (c) shows a single arteriovenous crossing which appears as two bifurcation points.
- Now, to differentiate between a junction that is bifurcation or arteriovenous crossing, the local variance-based method is used. A circular window of radius 11 is employed here that is made centered on the detected junction points (as shown in Fig. 10(d), where the junction points are shown in red with a circular window in green. The variance of pixels inside the window in green component of RGB image is captured. Since the vessel crossings do not contain pixels from just one type of vessel, i.e., it either contains pixels from (artery and vein) or (artery and capillary) or (vein and capillary), the captured variance will show a spike in variance if it is a cross-over and will have more variance than bifurcation points. This variance-based property is used here to characterize the vessel junctions as cross-overs or bifurcations.
- Once the bifurcations and crossing-overs have been differentiated, the cross-over points are

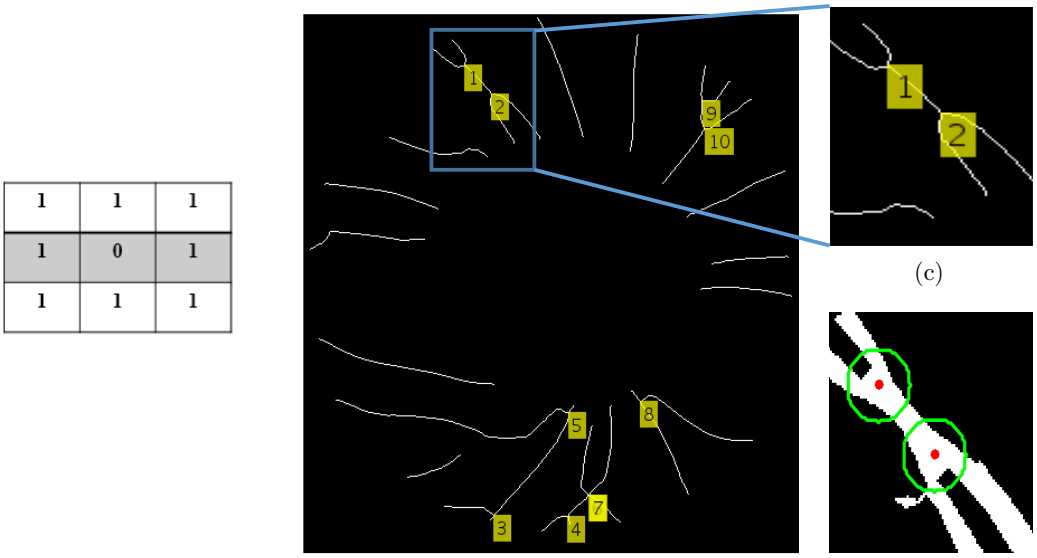


Fig. 10. Detection of vessel junctions. (a) 3×3 window used to detect the junctions, (b) Location of vessel junctions, (c) An arteriovenous crossing which fakes as two vessel bifurcations and (d) Circular window (in green) centered on junction points in (red).

eroded in original extracted vessel map. As a result, we will have an image that does not have vessel cross-overs.

It should be noted that for detection of cross-over points, different structures with varying radii were tested and the preliminary trials favored the choice of circular window with radius 11. After the cross-overs are identified and removed, the binary image now contains veins, arteries and small thin capillaries. However, due to unavailability of ground truth, the thin capillaries are not considered in next phase, i.e., feature extraction for vessel classification.

2.5. *Feature extraction for vessel classification*

Once the vessel tree is split into vessel subsegments, features are extracted from each candidate vessel segment. Each detected vessel is regarded as a sample for classification and represented by a feature vector containing several features. In previous work, we proposed method for vessel classification.[67] However, the research was tested on a small image dataset that does not contain any pathology. In another approach,[68] retinal vessels were classified using a small number of vessel segments from each dataset (major A/V pairs).

In this paper, 81 features are proposed for representation of blood vessels in retina, i.e., a single vessel sample is represented by 81 features. Although, the use of large number of visual representations for a pattern recognition problem increases the probability of accurate object classification, but it also increases the computations involved in extracting and categorizing those features. However, our proposed method includes the feature selection process, so any irrelevant features extracted will be removed before final vessel labeling. Let x_i be the candidate vessel sample considered for retinal vessel classification, where $i = 1, 2, \ldots, M$, and M being the total number of vessel samples. The number of vessel segments, i.e., i considered from each dataset is different. The true class label of x_i vessel sample is y_i, where y_i can only take on from two values, since the number of vessel classes is two, i.e., artery and vein. Denoted by $F = f_1, f_2, \ldots, f_N$, the feature matrix where $j = 1, 2, \ldots, N$ and N being the total number of features extracted from each vessel sample. So the dimensions of input data matrix X is $M \times N$, where each vessel sample x_i in X is represented by N distinct feature values. Y is the vector containing values of true class labels y_i. The features extracted for retinal vessel classification comprises of: (I) First-order statistical features characterizing the properties of histogram of vessel pixels in different color spaces. (II) Spatial

distribution of gradient magnitude representing the changes in vessel intensity with respect to fundus area pixels. (III) Features based on filter responses, where the filters borrowed from Leung and Malik,[69] Schmid,[70] Geusebroek *et al.*[71] These features are extracted from either vessel centerlines or complete vessel segments. Features extracted from vessel center-lines are effective due to the presence of light reflex in center of arteries and this light reflex in arteries makes them distinguishable from veins since veins are darker in intensity. Another reason for widely use of features from center-lines is the less number of pixels required in feature set which leads to less computations in form of pixel-processing as compared to features from complete vessel segments. RGB, CIE L*a*b, CMYK and YCbCr color spaces are exploited for extraction of features. All the features are concatenated into a single feature set, consisting of 81 features. It has been observed in literature that features obtained using different methods outperform single type of features because visual representations acquired using multiple techniques have the ability to capture various aspects of same object in image.[72,73] Details of proposed features are elaborated below and Table 1 tabulates the features which are used in this paper for retinal vessel recognition.

2.5.1. *Histogram-based first-order statistical features*

First-order statistical features are extracted using histogram of vessel center-lines. The vessel center-line pixel intensities are recorded in different color spaces and saved into a vector. Then the distribution of vessel center-line pixel intensities is analyzed by quantizing the total intensity range into 256 bins. The motivation behind extracting features from histogram-based statistical values is that they offer better results as compared to raw pixel values.[74] 21 features are extracted using statistical properties of intensity distribution as described in Table 1.

2.5.2. *Vessel Intensity Transition Features (VITF)*

VITF represents strength of change in intensity as the circular profile crosses the vessel segments, illustrated in Fig. 11 and this strength is expressed here by means of gradient magnitude. Gradient magnitude accurately depicts the power of peaks and valleys of intensity in an image. The cross-sectional vessel intensity transitions are investigated in depth here. As mentioned before that arteries are

Table 1. Details of feature set.

Feature no.	Feature description	Type
1–21 ($f_1 - f_{21}$)	Mean, Standard Deviation, and Entropy of vessel Center-line pixels. (Red and Green channel of RGB; Luminance and Chrominance channel of YCbCr color space, L and B channel of L*a*b and M of CMYK)	Histogram-based statistical features
22–49 ($f_{22} - f_{49}$)	Minimum and Maximum value of gradient magnitude, Kurtosis and Variance of gradient magnitude histogram. (Red and Green channel of RGB; Luminance and chrominance channel of YCbCr color space, L and B channel of L*a*b; and M of CMYK)	Gradient magnitude and spatial distribution of gradient magnitude
50–70 ($f_{50} - f_{70}$)	Minimum, Maximum and Variance of pixel intensities of complete vessel segment. (Red and Green channel of RGB; Luminance and chrominance channel of YCbCr, L and B channel of L*a*b, color space and M channel of CMYK)	Values directly calculated from complete vessel pixel intensities
71 (f_{71})	Ratio of center-line pixels intensity average to average of intensities in complete vessel segment (in Green channel of RGB)	Average of pixel intensity values taken from center-line pixels and complete vessel segments
72–81 ($f_{72} - f_{81}$)	Minimum and Maximum value of responses of filters from LM, Schmid and MR filter banks in red and green channels of RGB image	Values calculated from raw pixel values in filtered images

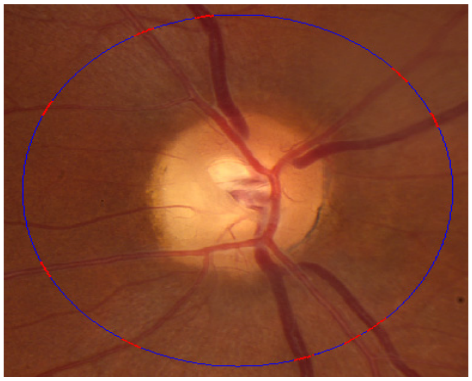


Fig. 11. Color retinal image with circular profile marked in blue and small curved portions in red.

brighter in intensity as compared to veins, so the intensity transitions captured for both classes (artery and vein) show significant difference. It should be noted that the curved profile is made to obtain features from all those vessel segments which are considered in RoI for vessel classification. For extraction of the transition features, a portion of circle cutting the vessel segments is taken and gradient magnitude of intensities along with that curved portion is calculated as shown in Fig. 11. The strength of gradient magnitude for vein class is larger due to abrupt changes in intensity of vein, whereas for artery class, they are lower because arteries are brighter as compared to veins so intensity transitions are not sharp as compared to fundus area pixels. Therefore, the gradient magnitude will be high for vein class as compared to artery class. This is characterized by considering the minimum and maximum value of gradient magnitude. Also, the spatial distribution of gradient magnitude is obtained and quantized into 15 bins. It is noticed that for vein class, the majority of values cluster at starting and ending bins in spatial distribution while for artery class, magnitude gradient values show membership to the bins which are in middle. Moreover, the gradient magnitude values for the artery class are evenly distributed among all bins as compared to vein class, so this behavior is captured by taking kurtosis and variance of gradient magnitude histogram. We further explain this property as follows: As the vein class shows wider blood vessels than artery class, the gradient magnitude is higher along its edges, but lower within its internal pixels. While for artery class, as the width of blood vessels is very narrow, the variation in gradient magnitude is less, so its spatial distribution of gradient magnitude is even. Different values of bins (varying from 10 to 30) are tested and it is observed that value of 15 bins capture most suitable characteristics of both vessel types. Based on these properties, 28 features are extracted in different color spaces, details are mentioned in second row of Table 1.

2.5.3. *Features from all pixels in vessel segment*

Next, 28 features are extracted using all pixel values in vessel segment, as described in third row of Table 1. The raw pixel values are considered for feature extraction in this category.

2.5.4. *Features from filter response of images*

Filters from Leung–Malik (LM) Filter Bank,[69] Schmid (S) Filter Bank[70] and MR Filter Bank[71] are employed here to extract features. The selected filters are convolved with red and green channel of RGB fundus image and the vessel center-line pixels of filter response images are considered for feature extraction.

Leung–Malik filter bank contains; 48 filters including first and second derivative of Gaussian filters at various orientations and scales, Laplacian of Gaussian filters and simple Gaussian filters. However, we have borrowed only one Gaussian filter which is at scale $= 2\sqrt{2}$. The selection of this filter is made through the preexperiments in which only one filter was seemed to enhance the retinal vessels.

Schmid filter bank contains 13 rotationally symmetric filters of the form given in Eq. (2).

$$F(r,\sigma,\tau) = F_o(\sigma,\tau) + \cos\left(\frac{\pi\tau r}{\sigma}\right) e^{\frac{r^2}{-2\sigma^2}}, \qquad (2)$$

where (σ,τ) takes the values (2,1), (4,1), (4,2), (6,1), (6,2), (6,3), (8,1), (8,2), (8,3), (10,1), (10,2), (10,3) and (10,4). In this paper, we have used two filters from Schmid filter bank, with (σ,τ) having values of (2,1) and (10,4). The third filter bank used is MR filter bank, which basically consists of 38 filters. From this filter bank, we have borrowed two isotropic filters, i.e., Gaussian filter and Laplacian of Gaussian filter. Ten features are extracted using the filtered images, as shown in fifth row of Table 1.

2.6. *Width estimation of vessels*

The width of classified vessel segments is calculated here using 2D Euclidean distance transform.[75] When the distance transform of a complement of binary image is taken, the resultant image looks like a gray-level image, but in actual distance-transformed image represents distance of the respective pixel to the nearest nonzero pixel. For calculation of width, first the complement of segmented binary vessel image is taken (say I_1) shown in Fig. 12(b). 2D Euclidian distance transform is applied on I_1 in Fig. 12(b), giving I_D as a resultant image shown in Fig. 12(c) with its enlarged version. Then the center-line vessel map of original binary vessel network I_B is obtained, as shown in Fig. 12(d) and multiplied with I_D to acquire distance map value for center-line pixels. This distance map image (shown in Fig. 12(e)) is finally multiplied by two to get vessel width.

2.7. *Calculation of AVR*

After we have acquired the width of vessels, AVR is calculated. The parameters Central Retinal Arterial Equivalent (CRAE) and Central Retina Venous Equivalent (CRVE) are determined using Parr–Hubbard formulas.[64,65] It should be noted that the width of single vessel does not remain constant for whole vessel and varies according to the pixels as shown in Fig. 12(e). Therefore, mean of width of a vessel segment is taken and collected in separate vector; "Arteriole" and "Venule", respectively, depending upon the class of label of vessel generated previously. These two vectors show the mean width of respective vessels. Equations (3) and (4) show the formula for calculating CRAE and CRVE, respectively.

$$\text{CRAE} = \sqrt{(0.87W_a^2 + 1.01W_b^2 - 0.22W_aW_b - 10.73)}, \tag{3}$$

where W_b and W_a, is the median and the value occurring immediately before the median in vector "Arteriole", respectively.

$$\text{CRVE} = \sqrt{(0.72W_a^2 + 0.91W_b^2 + 450.02)}. \tag{4}$$

Likewise for CRVE, W_b and W_a is the median and the value occurring immediately before the median in vector "Venule", respectively. AVR is calculated as given in Eq. (5).

$$\text{AVR} = \frac{\text{CRAE}}{\text{CRVE}}. \tag{5}$$

3. Feature Selection and Classification of Vessels

In this section, the method proposed for feature selection and vessel classification is explained. To select the significant features, the features are first

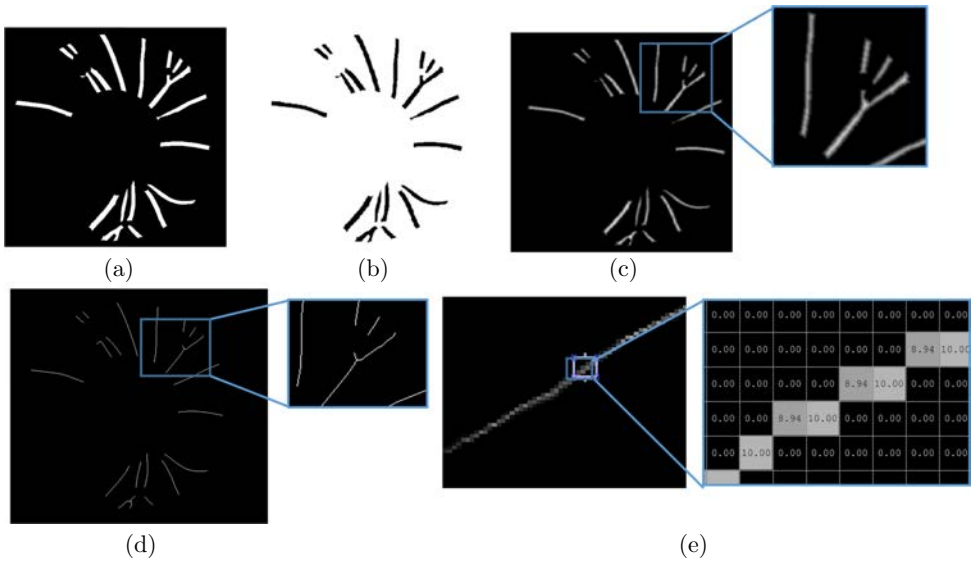


Fig. 12. Vessel width computation. (a) Binary vessel map, (b) Complement of binary vessel map, (c) Distance map of complement binary map, (d) Center-line vessels and (e) Product of (c) and (d).

ranked using two feature ranking strategies and then the classification accuracy of three classifiers is used as a threshold to select an optimal feature subset from ranked feature list. The feature subsets selected depending on accuracy of each classifier are then fused to make a single feature subset and will be used by hybrid labeling method for retinal vessel classification. This process is elaborated in detail in sections to follow. We will first discuss the feature selection process here followed by vessel classification scheme.

3.1. *Ranking of features and selection of optimal feature subset*

The motivation for including a feature selection module is that the extracted features may contain some redundant data that leads to overfitting of prediction model and consequently reduction in prediction accuracy. The selection of features ensures the inclusion of only those features that are actually useful for classification and subsequently decreases the computational complexity.[76] In this paper, we use two feature-ranking techniques to rank the features but a novel approach is followed for selecting features from ranked feature list. The system works by first ranking the features according to two feature ranking methods, i.e., Pearson Correlation Coefficient[58] and Relief-F method[59] and then selection of optimal features. Generally, after the features have been ranked and arranged according to their ranks, a specific threshold is used to select a certain number of top-ranked features. This threshold is usually user-defined, however, as pointed out in Refs. 77 and 78, the correct way to ensure the selection of optimal combination of top-ranked features is by evaluating the classification performance of different combinations of top-ranked features. Therefore, the selection of significant features from a ranked feature list cannot be carried out using a fixed threshold because we do not have a prior knowledge regarding the performance of different number of ranked features. In this paper, we use the maximum classification accuracy of classifiers as the threshold to select the features. A particular number of top-ranked features that yield maximum accuracy on classifier is selected. In this paper, we employed three classifiers for selection of optimal number of top-ranked features. The optimal feature subsets selected using three classifiers are combined and used by proposed hybrid classification technique for vessel classification.

Class label y_i and feature value f_j of every sample x_i is given as an input to feature ranking strategy. Both these techniques evaluate correlation of each feature with the class label using some criteria[79,80] and a rank is generated for each feature. The features are then arranged in descending order of their ranks, i.e., $F^\circ = \{f_{r_1}, f_{r_2}, f_{r_3}, \ldots, f_{r_N}\}$, where f_{r_1} and f_{r_N} denotes the features with highest and lowest rank, respectively. For selection of feature subset, the ranked features F° are given as an input by constructing "n" feature subsets, in which first feature subset is initialized by incorporating only the highest-ranked feature, the second subset is constructed by adding second top-ranked feature in the first subset and this process is repeated until the last feature subset contains all ranked features. By applying different combinations of highly ranked features to classifiers, different predictive models are generated which have different accuracies. The optimal predictive model is the one with maximum classification accuracy and the feature subset corresponding to this predictive model is selected. This feature selection process is shown in Fig. 13. As illustrated in Fig. 13, different combinations of ranked features are applied to classifier, and the feature subset, f_S^T, corresponding to maximum classification accuracy, T, is selected. Now, since we have employed three supervised classifiers, so three optimal feature subsets are obtained. The number of features in the optimal subsets is not specific. This hybrid approach allows us to select different feature combinations from ranked feature list. The ranked optimal feature subsets obtained using Pearson Correlation Coefficient method on k-NN, SVM and Naïve Bayes classifier are denoted by f_{p_1}, f_{p_2} and f_{p_3}, respectively, and those obtained using Relief-F method on k-NN, SVM and Naïve Bayes classifier are denoted by f_{p_1}, f_{p_2} and f_{p_3}, respectively. Finally, the union of those optimal feature subsets is taken as given in Eqs. (6) and (7), and the resulting feature subsets, F_P and F_R are used by proposed hybrid classification scheme for retinal vessel labelling. The feature ranking approaches used are

$$F_P = f_{p_1} \cup f_{p_2} \cup f_{p_3}, \tag{6}$$

$$F_R = f_{r_1} \cup f_{r_2} \cup f_{r_3}. \tag{7}$$

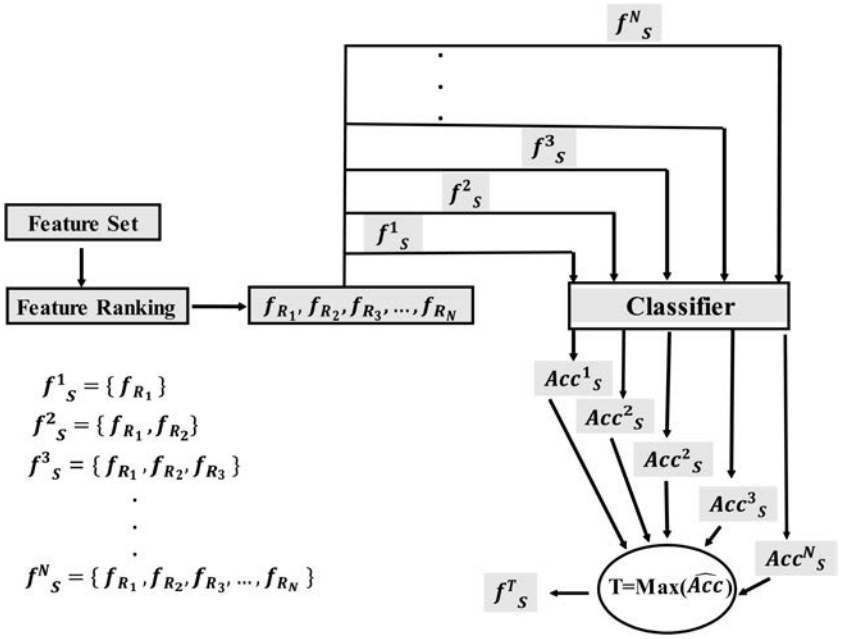


Fig. 13. Complete framework for optimal ranked features selection, where f_S^N represenets the total number of feature subsets acquired from ranked feature list, Acc_S^N denotes the classification accuracy corresponding to feature subset f_S^N, T represents the maximum accuracy obtained using a feature subset f_S^T.

3.1.1. *Pearson Correlation Coefficient method*

The Pearson Correlation Coefficient (PCC) method ranks features by calculating linear correlation between individual features and class labels.[58] In this paper, we use PCC method to obtain rank of features. This PCC method finds correlation p_i for relevance assessment of the feature f_j with the corresponding class label y_i and as an output, a correlation score for each individual feature is generated. PCC of a vessel sample x_i (where $x_i \in X$) and class label y_i (where $y_i \in Y$) is calculated as given in Eq. (8), where cov is covariance and σ is variance.

$$p(x_i, y_i) = \frac{\mathrm{cov}(x_i, y_i)}{\sqrt{\sigma(x_i)\sigma(y_i)}}. \tag{8}$$

3.1.2. *Relief-F algorithm*

The second feature ranking algorithm that is used here is Relief-F algorithm.[59] In Relief-F algorithm, each feature gets a weight depending upon its strength for distinguishing between those opposite class samples that are near to each other and difficult to differentiate. The feature rank is calculated by taking a data point at random and considering the k-nearest neighbors of that data point.[59] The k-nearest neighbors are taken from both the classes and by considering their contribution, the strength of feature is analyzed. In order to find the optimal value of k for this feature ranking technique, we analyzed the weights of features by varying the value of k. Here, maximum value of K is taken as 75% of total instances because according to theory of Relief-F concept,[59] if the value of K is taken too small, then estimates would be difficult to generalize on highly varied data whereas if the value of K is equal to number of instances, then significance of relevant features will be deteriorated.

3.2. *Details of classifiers used for feature selection after ranking of features*

After the features are ranked using the above mentioned strategies, significant features are selected by application of ranked features on three classifiers (k-NN, SVM and Naïve Bayes). The same three classifiers are used in decision fusion framework for recognition of retinal vessels. The purpose to use the same classifiers for vessel classification, is to study the effect of single classifiers performance with the ones obtained when they are combined. The classifiers used are detailed in what follows.

3.2.1. *k-NN classifier*

k-NN is one of the simplest classifiers used for supervised classification.[81] It searches for closest k samples from complete dataset by calculating distance between training and test instances. For example, in our case, when an unknown vessel sample, say x_i comes, the labels of k-nearest neighors of x_i are analyzed and then x_i is assigned a class (either "artery" or "vein") depending upon the label of majority neighbors. The distance between the test sample and all nearest neighbors is calculated. "Euclidean distance" is chosen to calculate the distance between sample and neighbors. Equation (9) shows the Euclidean distance calculated between the vessel sample and its nearest neighbors, where x_i is the test sample and x_{mb} represent the nearest neighboring samples with subscript b showing the total number of nearest neighbors. In order to obtain the optimal value of k for our vessel classification task, different k values such as 1, 3, 5, 7 and 9 are tested.

$$d(x_i, x_m) = \sqrt{\sum_{b=1}^{q} (x_i - x_{m_b})^2}. \tag{9}$$

3.2.2. *SVM classifier*

SVM is a supervised machine learning method that separates different classes in testing data by an optimal hyperplane[82] and the advantage of using SVM in classification lies in its ability to identify a nonlinear separation between data-points of different classes. This is carried out by using different "separation" functions called kernels. The input sample vector x_i represented by N feature values is mapped to a new feature space φ with higher demensions and an optimal hyperplane is constructed using the kernel $\mathrm{Kl}(x_i, x_m)$, shown in Eq. (10), where x_i and x_m are two feature input vectors. In our case, the dataset is tested using SVM with different kernel functions, i.e., Polynomial ($\mathrm{Kl}_{\mathrm{po}}$) and Radial Basis Function (RBF) ($\mathrm{Kl}_{\mathrm{Ga}}$), and the function that provides optimal results is selected. For polynomial kernel, the degree is varied from 1 to 3 and for RBF kernel, the scaling factor is tested using values from 1 to 9. The representations for both these kernel functions are given in Eqs. (11) and (12).[83]

$$\mathrm{Kl}(x_i, x_m) = \langle \varphi(x_i).\varphi(x_m) \rangle, \tag{10}$$

$$\mathrm{Kl}_{\rho o}(x_i, x_m) = (x_i.x_m + 1)^{\rho}, \tag{11}$$

where ρ is degree of polynomial (11)

$$\mathrm{Kl}_{\mathrm{Ga}}(x_i, x_m) = e^{\frac{||x_i - x_m||^2}{2\sigma_G}}. \tag{12}$$

where σ_G is Gaussian sigma.

3.2.3. *Naïve Bayes classifier*

Naive Bayes classifier is a widely used probabilistic classifier that classifies the data based on Bayes' Theorem.[84] The input data matrix X and vector of labels Y are fed to classifier, and according to Bayes theorem, the classification aim is to achieve maximum probability $P(Y|X)$, as shown in Eq. (13). The strength of Naïve Bayes is its simplicity and efficiency with which it classifies the data. Naïve Bayes classifier is implemented using different distribution functions such as kernel and normal. The classifier is tested with both distribution functions and model with highest validation accuracy is selected.

$$P(Y_v|X) = \frac{P(X|Y_v)P(Y_v)}{P(X)}. \tag{13}$$

3.2.4. *Proposed hybrid classification scheme*

The proposed hybrid classification technique is a fusion of decisions generated by k-NN, SVM and Naïve Bayes classifiers for vessel classification. The proposed scheme works in a way such that the labels given by each of three classifiers are acquired and a final label is assigned to the vessel sample by a classifier fusion technique known as majority voting.[85] The main motivation for combining the decisions of the supervised machine learning methods is the strategy used for decisions combination, which reflects the local competency of individual learning methods.[86] Moreover, the samples misclassified by individual classifiers may not necessarily overlap.[87] Majority voting is one the popular classifier fusion techniques in which the single labels produced by individual classifiers are counted and the sample is given the label with majority votes.[87] It is an efficient classifier fusion scheme since it does not require any other information except for single class labels generated by individual classifiers. The optimal feature set selected by the feature selection process is used by proposed hybrid classification method for vessel differentiation.

4. Experimental Results

Through experimental results, we aim to investigate if the proposed feature selection and classifier fusion technique improves the retinal vessel classification task and subsequent AVR calculation. We first give a description of datasets used in this research. Then, we elaborate the procedure for choosing the optimal parameters of classifiers. Afterwards, we illustrate the feature ranks obtained from two different strategies and features selected after application of ranked features on classifiers. Then, we show the results of applying selected features on proposed decision combination framework followed by AVR calculation results. Computer program in this work is implemented using computer with 1.80 GHz processor and 4.0 GB RAM. Commercial software MATLAB is used for implementation purpose.

4.1. *Specifications of datasets*

The methodology is evaluated on three databases: a database collected from Armed Forces Institute of Ophthalmology (AFIO), Pakistan, and two public labeled databases, i.e., INSPIRE-AVR[33] and VICAVR database.[31] In this research, only those vessel segments are used for retinal classification and quantification purpose whose labels are provided in ground truth.

The local database contains 44 retinal images in JPEG format with dimensions 1504×1000, including 11 images containing pathological structures like hard exudates, cotton wool spots, hemorrhages, arteriosclerosis, vessel tortuosity, focal arteriolar narrowing and OD blurring. True vessel labels and AVRs are acquired by an expert ophthalmologist that will be are considered as a ground truth. INSPIRE-AVR is a publically available database containing high-resolution 40 OD centered healthy retinal images, acquired at the University of Iowa Hospitals and Clinics. These images are of size 2392×2048 and available in JPEG format. AVR values estimated by two observers are provided with INSPIRE-AVR database, to be used for comparison and the vessel labels are acquired from our ophthalmologist. Third database used is VICAVR that contains OD-centered 58 images of size 768×576. The artery-vein labels and vessel caliber for VICAVR database, obtained from three human experts are available with the dataset. Table 2 describes the complete specifications of datasets used in this study. The images in local dataset are marked by ophthalmologist on the basis of visual appearance; therefore, underlying causes of different abnormal structures are not exploited here. Additionally, detection and diagnosis of the other retinal pathologies occurring independently or associated with HR are beyond the scope of this research. In this paper, we have evaluated the results only for vessel classification and AVR computation. The unclassified vessels are not included in evaluation.

4.2. *Parameter tuning of classifiers*

In our research, classification accuracy of three classifiers (k-NN, SVM and Naïve Bayes) is used as a threshold to select the optimal top-ranked features. Before giving ranked features as an input to classifier, for feature selection, the parameters of classifier are tuned. This is done in "parameter tuning phase", where the dataset is divided into two parts, training set (70% of data) and validation set (30% of data). The classifier is tested with different parameters using training data and then accuracy is evaluated on validation set. Complete ranked feature set of 81 features is given as input to classifier for tuning of parameters. The parameters showing maximum accuracy on validation set are selected. All three supervised classifiers are trained once using training data and then tested on validation set to acquire optimal parameters. The values of optimal parameters of classifiers are mentioned in Table 3.

Once the optimal parameters are acquired, three classifier models (k-NN, SVM and Naïve Bayes) are refit again to entire dataset using 10-fold cross validation and classification accuracy of these three classifiers is used to select optimal number of ranked features from ranked feature list. In 10-fold cross-validation, data is divided into 10 subsets, out of which nine are retained for training and one is used for testing. The samples which are included in each fold are randomly selected. Each fold is iteratively tested and the rest of folds are kept for training. The 10-fold cross-validation is conducted for 10 times, and the samples which are included in each 10-fold cross-validation are randomly selected differently. Different combinations of ranked feature subsets are given as an input to classifier and the subset leading to maximum classification accuracy is selected. The performance metrics calculated for evaluation of ranked feature subsets on classifiers are, accuracy,

Table 2. Specifications of datasets.

Datasets	Number and size of images	Number of vessel segments for classification	Ground truth and details of pathologies
Local dataset	44 images of dimensions (1504×1000)	356 vessel samples (195 vein vessels, 161 artery segments)	Corresponding AVR values for 44 images estimated by expert ophthalmologist Images include: 20 Non-HR images 11 images with Grade-I HR including (4 images with vessel sheathing near OD, 2 images with hard exudates and hemorrhages, 1 image with branch retinal vein occlusion) 8 images with Grade-II HR including (1 image with vessel sheathing near OD, 1 image with cotton wool spots and hemorrhages) 3 images with Grade-III HR including (2 images with cotton wool spots and hemorrhages) 2 images with Grade-IV HR including (2 images with cotton wool spots, hemorrhages and OD swelling)
INSPIRE-AVR dataset	40 retinal images of dimensions (2392×2048)	410 vessel samples (201 artery segments, 209 vein segments)	Corresponding AVR values for 40 images estimated by two Observers
VICAVR database	58 retinal images of dimensions (768×576)	476 vessel samples (244 vein vessels, 232 artery segments)	Artery/Vein labels and AVR values for 40 images estimated by three Observers

Table 3. Parameters configuration for different classifiers.

Classifiers	Classifier parameters selection using features ranked by both methods
k-NN	Nearest neighbors $k = 3$
SVM	"RBF" kernel with scaling factor 7
Naïve Bayes	"kernel" distribution function

sensitivity and specificity. However, accuracy metric is used as a stopping criteria to select the number of top-ranked features from ranked feature list or in other words, as a threshold to select the optimal feature subset. Sensitivity the is true positive rate (positives) and specificity is true negative rate (negatives). The feature selection procedure is illustrated in Fig. 13. The evaluation parameters are calculated using Eqs. (14)–(16), respectively.

$$\text{Sensivity} = \frac{T_P}{(T_P + F_N)}, \tag{14}$$

$$\text{Specificity} = \frac{T_N}{(T_N + F_P)}, \tag{15}$$

$$\text{Accuracy} = \frac{(T_P + T_N)}{(T_P + T_N + F_P + F_N)}. \tag{16}$$

4.3. *Performance of feature raking methods and Selection of optimal features*

In order to rank the features using Relief-F method, the value of K, i.e., number of nearest neighbors is determined. We analyzed the weights of features with varying the number of nearest neighbors, i.e., $K = 1$ to 250, 1 to 287 and 1 to 333, for local, INSPIRE-AVR and VICAVR datasets, respectively. The optimal value of K for which the weights of features become stable is 196, 132 and 245 for local, INSPIRE-AVR, and VICAVR dataset, respectively. Figure 14 shows the sample plots of first 24 features from local database. The ranking of features depends on the weights, and Fig. 14 shows the weights of features are varying with increasing the number of K-neighbors. It can be observed in Fig. 14(a) that the feature 1 has a significant weights difference with other 5 features (i.e., from

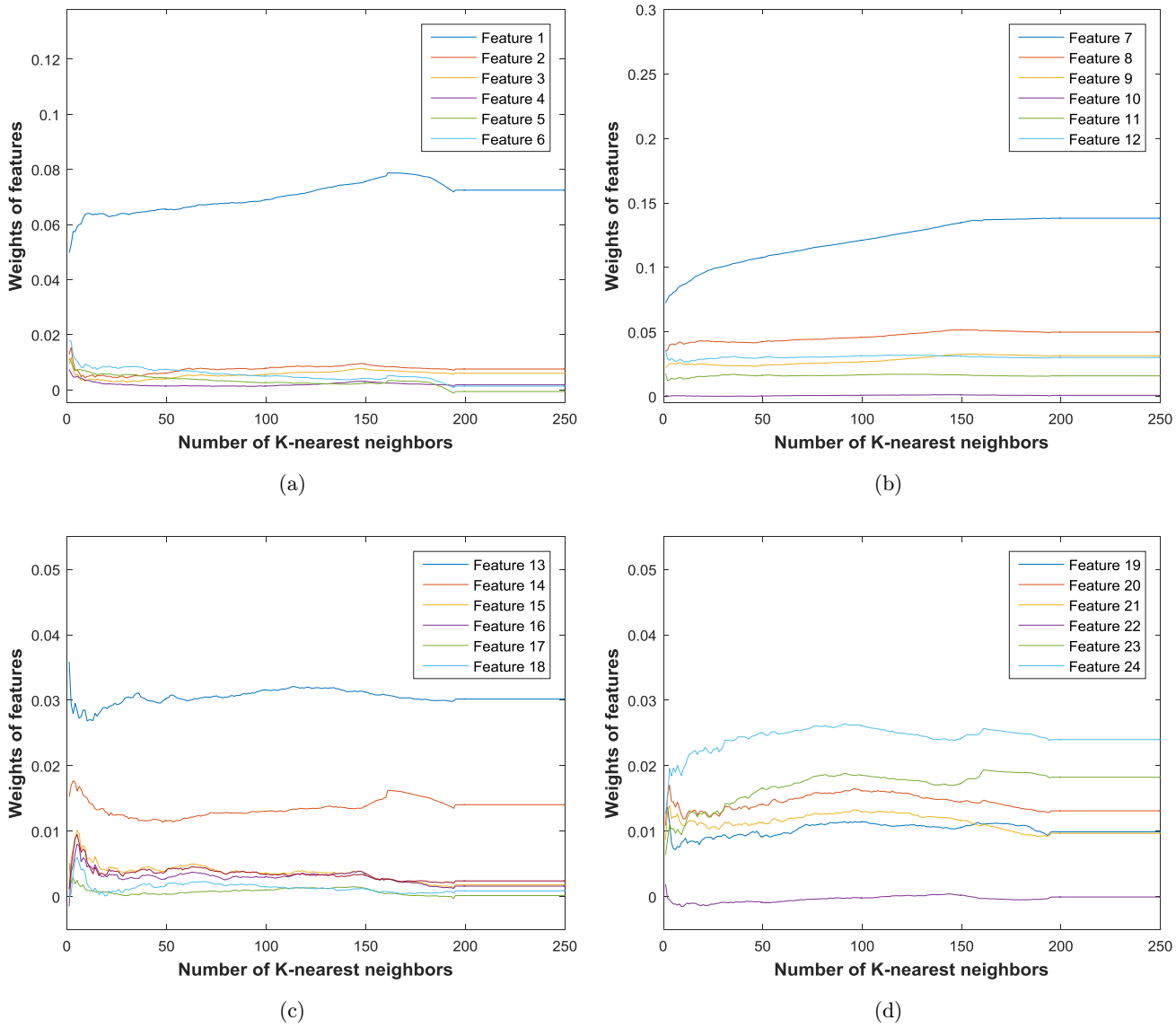


Fig. 14. Weights of first 24 features varying as number of neighbors are increased from 1 to 250 by Relief-F algorithm (for local database). (a) Values of feature weights from 1 to 6, (b) Varying weights of features from 7 to 12, (c) Varying weights of features from 13 to 18 and (d) Varying weights of features from 19 to 24.

2 to 5) and the features from 2 to 6 have negligible weight difference. In Fig. 14(b), it is observed that the weights of feature 7 are increasing with an increase in number of neighbors. It can also be observed that as K approaches 196, the feature weights become stable. This situation motivates us to select $K = 196$ for local database, since adding more neighbors is not contributing towards the better modeling of data.

For selection of top-ranked features, different subsets of features acquired from ranked feature lists are applied on k-NN, SVM and Naïve Bayes classifier using 10-fold cross-validation. This lead to generation of 81 predictive models with different accuracy, sensitivity and specificity metrics, as illustrated in Fig. 15 with upper and lower curves showing performance metrics for local and INSPIRE-AVR dataset, respectively. The feature subset that maximizes the classification accuracy on each classifier is selected. For a better visualization and comparison purpose, we concatenated the performance curves of both datasets on same x-axis in Fig. 15. From these illustrations, the relation of classifier with certain combinations of ranked features can be observed. Each classifier's response for each feature subset is different, which indicates the

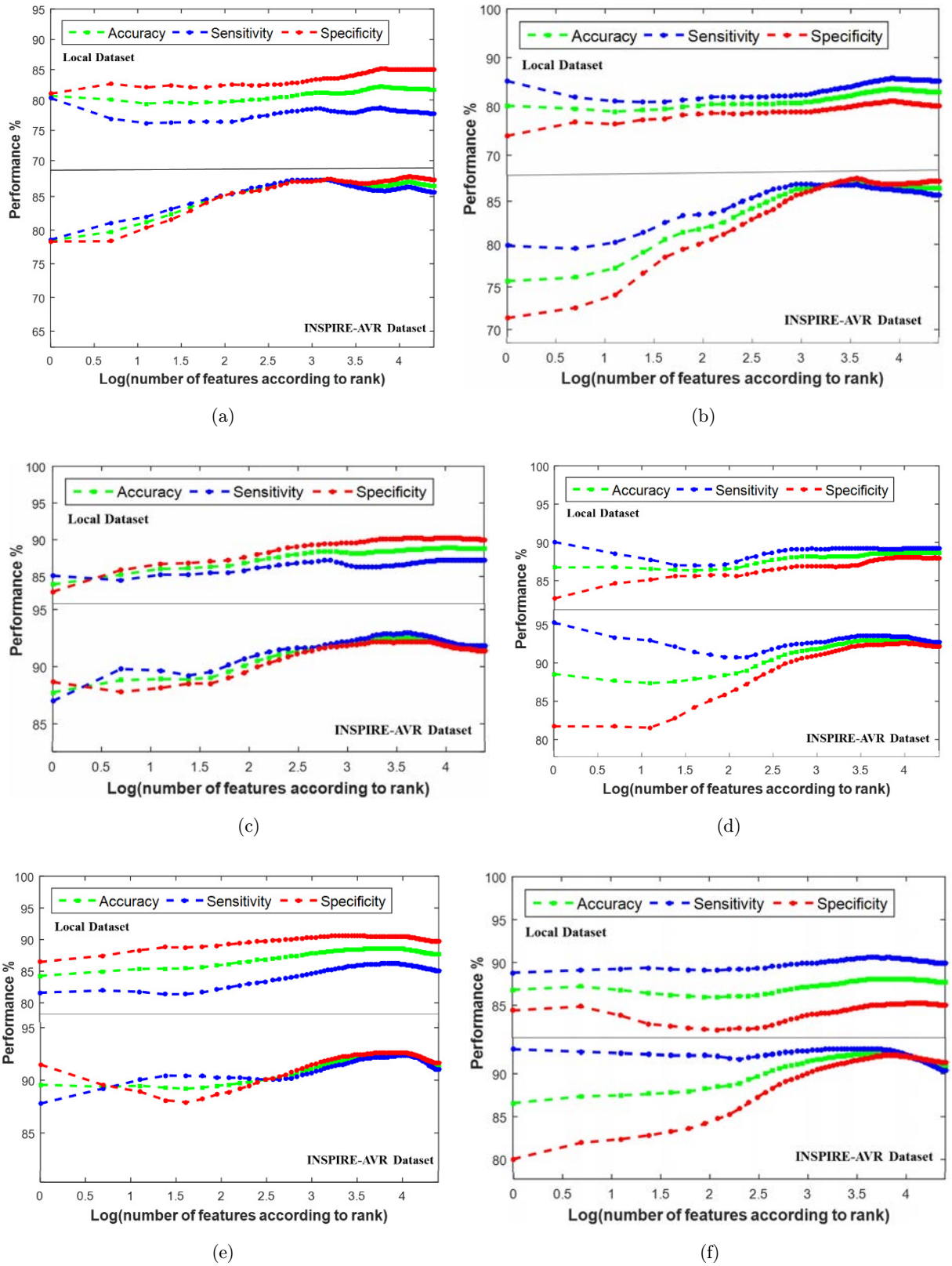
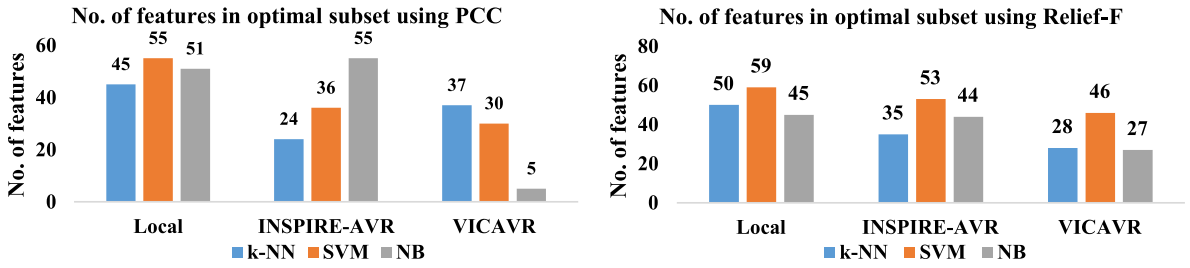


Fig. 15. Performance metrics (Green, Blue and Red curves depicting Accuracy, Sensitivity and Specificity, respectively) for different-ranked feature subsets, upper curves: Local dataset and lower curves: INSPIRE-AVR dataset. (a), (c) and (e) Classification performance of k-NN, SVM and Naïve Bayes classifier on features ranked by PCC and (b), (d) and (f) Classification performance of k-NN, SVM and Naïve Bayes classifier on features ranked by Relief-F method.

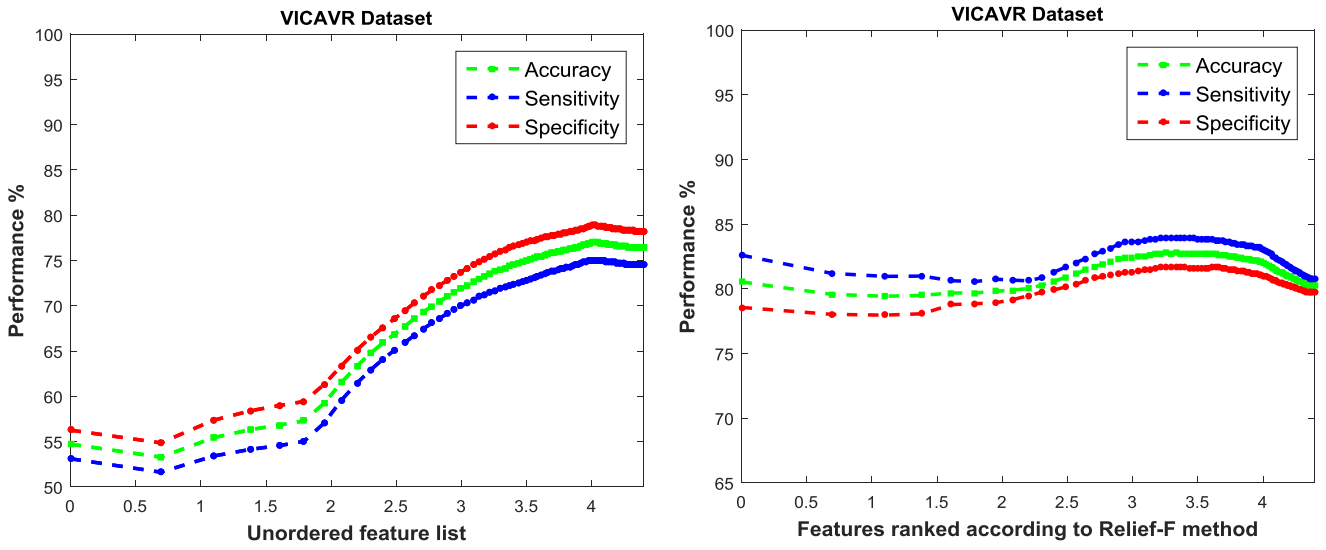


(a) (b)

Fig. 16. Number of features in optimal feature subsets acquired from two ranking schemes with classification accuracy of three classifiers. (a) Number of features in optimal feature subsets using PCC method and (b) Number of features in optimal feature subsets using Relief-F method.

importance of features combination on the class label outcomes. Use of different classifiers for selection of optimal number of features, resulted in selection of feature subsets containing different number of features. For example, as shown in Fig. 16(a), the feature subset that led to maximum classification accuracy on k-NN, SVM and Naïve Bayes classifier for local dataset using PCC, consists of different number of features, i.e., 45, 55 and 51, respectively. These feature subsets will be combined using union operation to obtain a final feature subset.

In terms of number of features, no specific pattern is seen among the number of features in optimal subsets in Fig. 16. However, in majority cases, SVM requires the largest number of features to reach the maximum classification accuracy. In all cases, the value of "threshold" (maximum classification accuracy), is highest for INSPIRE-AVR database. An obvious explanation for this observation is the greater quality of fundus images in INSPIRE-AVR dataset. From Fig. 15, it is also concluded that overall no significant difference is seen between the performance of classifiers on features ranked either using PCC or Relief-F method. We also tested if the use of feature ranking is beneficial by analyzing the effect of feature subsets obtained from; (a) Ranked features list using Relief-F method, (b) Raw feature list without using ranking algorithm, on k-NN classifier for VICAVR database. From Fig. 17, it is found that feature ranking strategies have actually contributed in increasing the classification



(a) (b)

Fig. 17. Comparison of classification accuracy achieved for raw feature subsets (unordered) and ordered feature list using Relief-F method on k-NN classifier for VICAVR dataset. (a) Classification performance of feature subsets from unordered feature list and (b) Performance attained using feature subsets from ranked lists acquired through Relief-F method.

accuracy. Note that the average accuracy and peak accuracy attained from ranked list is much higher as compared to accuracy illustrated using unordered feature list.

4.3.1. *Optimal feature subset selection*

Since we use classification accuracy generated by three classifiers as a criteria to select optimal feature subset, so three different feature subsets will be selected for each dataset. As it can be seen in Fig. 16, the number of features showing maximum classification accuracy on each classifier is different, even for the same dataset. So, in order to have a single feature subset that can be used by proposed hybrid classification approach, we combine the feature subsets acquired using different classifiers. The number of features in each subset resulted after talking union are shown in Fig. 18. The exact features in selected subsets are shown in Table 4. Majority of features in feature subset selected using PCC with preevaluations by three classifiers are also present in feature subset acquired using Relief-F method. This indicates the similarity in ranking of two different strategies.

4.4. *Retinal vessel classification using proposed hybrid classification scheme with optimal feature subset*

The proposed hybrid classification scheme combines the labels generated by three classifiers, i.e., k-NN, SVM and Naïve Bayes. This decision-combination is chosen for vessel classification because it represents the joint strength of objective function of multiple classifiers. In majority voting, the votes given by each classifier for a certain sample are counted and the class with maximum votes is assigned to the sample. For example, if SVM and Naïve Bayes assigns "Artery" to a sample whereas k-NN assigns "Vein", the final label will be given as "Artery", since the "Artery" class has two votes. After assigning the labels using proposed hybrid classification method, accuracy, sensitivity and specificity is calculated by comparing the final labels with ground truth. The classification performance achieved using proposed decision fusion framework with optimal feature subsets (F_P) and (F_R), acquired using PCC and Relief-F method, is given in Tables 5 and 6, respectively. Overall, hybrid classification scheme has caused to increase the vessel classification accuracy, both for feature subsets acquired from PCC and Relief-F ranking list. We have illustrated the increments in classification accuracies obtained with proposed classification technique with those acquired single classifiers associated with averaged classification accuracy on same feature subsets, as shown in Figs. 19 (using PCC) and 20 (using Relief-F). For all datasets, the increase in vessel classification performance show the improvement induced by the multi-classifier decision combination.

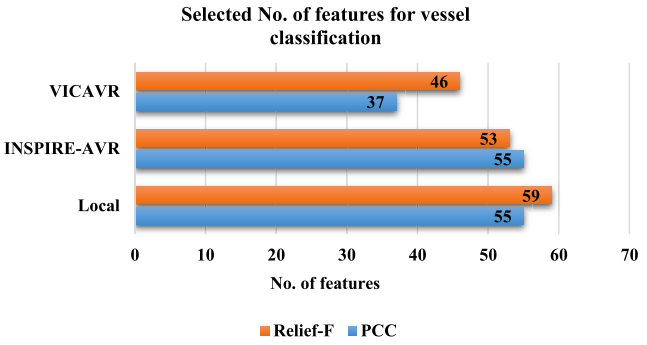


Fig. 18. Feature subsets selected for vessel classification, using two different ranking methods for VICAVR, INSPIRE-AVR and Local database.

Confusion matrix attained using the performance measures are shown in Fig. 21. Confusion matrix is a representation of overall classification performance, and reflection of the fraction of two classes being correctly classified or misclassified. These illustrations validates the ability of proposed system in classifying vessels. Here, we represent the class of artery and vein sample with "1" and "0", respectively. Therefore, the top two entities on the left diagonal of confusion matrix (in green color) are True Negative (TN) and True Positive (TP), respectively. In our case, TP and TN, is the fraction of vessels being correctly classified as, arteries and veins, respectively, by the proposed method. The other two entities in the adjacent diagonal (in peach color) represent the False Negative (FN) and False Positive (FP), respectively. We represent FP and FN, as the fraction of veins being misclassified as arteries and fraction of arteries being misclassified as veins, respectively. The first two entities in third row of confusion matrix represent the Sensitivity and Specificity, respectively. In our case, Sensitivity represents the percentage of TP (vessels correctly classified as arteries) and Specificity represents the percentage of TN (vessels correctly classified as

Table 4. Optimal feature sets obtained via fusion of features using PCC (F_P) and Relief-F (F_R) ranking.

Datasets	F_P	F_R
Local database	{55, 7, 25, 31, 53, 8, 69, 1, 78, 51, 9, 45, 26, 49, 70, 71, 12, 64, 24, 68, 48, 54, 27, 47, 77, 23, 44, 60, 66, 30, 11, 50, 67, 72, 20, 61, 13, 21, 29, 57, 59, 73, 19, 2, 52, 36, 42, 35, 3, 75, 63, 37, 34, 39, 76}	{7, 55, 25, 31, 53, 1, 51, 8, 69, 49, 78, 26, 45, 9, 12, 70, 47, 27, 24, 64, 54, 50, 30, 23, 66, 60, 71, 68, 11, 48, 61, 13, 44, 77, 67, 20, 57, 59, 72, 19, 21, 29, 2, 42, 36, 39, 37, 63, 3, 73, 52, 35, 38, 75, 58, 18, 65, 40, 14}
INSPIRE-AVR database	{25, 39, 51, 1, 55, 67, 38, 50, 7, 31, 53, 37, 49, 66, 69, 78, 54, 65, 70, 52, 71, 2, 77, 73, 56, 72, 40, 26, 3, 8, 74, 27, 32, 9, 68, 33, 13, 59, 58, 6, 57, 75, 30, 12, 5, 14, 63, 18, 62, 19, 15, 29, 17, 61, 76}	{1, 39, 51, 25, 50, 38, 31, 67, 55, 69, 78, 7, 70, 37, 49, 73, 53, 66, 52, 54, 65, 71, 77, 72, 5, 6, 22, 74, 75, 2, 40, 11, 13, 3, 23, 56, 29, 30, 35, 36, 12, 42, 57, 58, 16, 4, 59, 10, 68, 28, 26, 34, 61}
VICAVR database	{7, 69, 78, 12, 66, 70, 53, 11, 71, 55, 67, 77, 72, 45, 25, 31, 8, 59, 1, 75, 57, 13, 47, 73, 76, 61, 63, 54, 65, 19, 17, 18, 9, 58, 30, 64, 68}	{7, 12, 11, 66, 53, 8, 55, 67, 69, 71, 78, 77, 70, 72, 25, 45, 59, 31, 1, 57, 13, 75, 9, 73, 61, 63, 19, 54, 65, 18, 17, 30, 76}

Table 5. Performance of hybrid classification for feature subsets acquired using PCC method.

		Hybrid classification approach		
Datasets	Number of features in selected subset	Accuracy (%)	Sensitivity (%)	Specificity (%)
Local dataset	55	90.17	91.45	89.22
INSPIRE-AVR dataset	55	93.41	95.02	91.87
VICAVR dataset	37	87.82	85.08	90.79

Table 6. Performance of hybrid classification for feature subsets acquired using relief-F method.

		Hybrid classification approach		
Datasets	Number of features in selected subset	Accuracy (%)	Sensitivity (%)	Specificity (%)
Local dataset	59	90.45	90.45	90.45
INSPIRE-AVR dataset	53	93.90	95.52	92.34
VICAVR dataset	33	85.92	83.95	87.98

veins). The last entity in the third row on right side represents the overall classification accuracy. Here, it represents the percentage of all vessels being correctly classified. The other two entities in confusion matrix on extreme right side of first and second row are Positive Predictive Value (PPV) and Negative Predictive Value (NPV), respectively. PPV and NPV represent the percentage of TP and TN among all data samples of two classes, respectively. Here, PPV and NPV represent the fraction of arteries and veins, correctly classified among all artery and vein samples, respectively, in data.

The highest vessel classification accuracy is observed for INSPIRE-AVR dataset with 93.9% correct rate using Relief-F ranking and proposed hybrid classification method. An interesting observation is the error rate for classification of artery samples in local dataset, with 6.2% and 5.3% misclassification that is highest among all datasets. An obvious explanation for this result is the presence of pathological structures in local dataset images that deteriorate the appearance of arteries and thus make their classification more challenging. On the other hand, both for INSPIRE-AVR and VICAVR dataset, the misclassification rate for veins is more high as compared to arteries.

Figure 22 shows an example of two fundus images, selected from local database (Fig. 22(a)) with

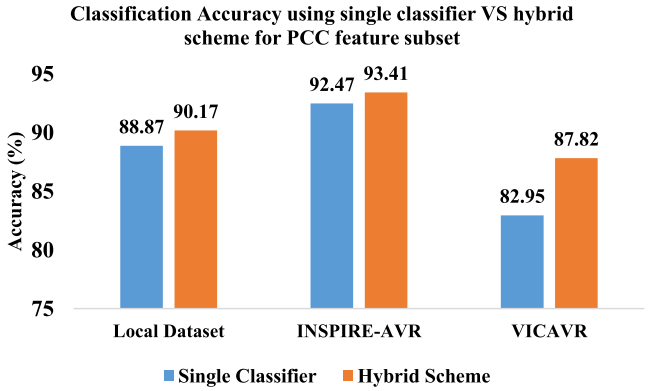


Fig. 19. Comparison of classification accuracy obtained using single classifier (shown in blue) and proposed hybrid classification (shown in orange), for feature subset from PCC ranking.

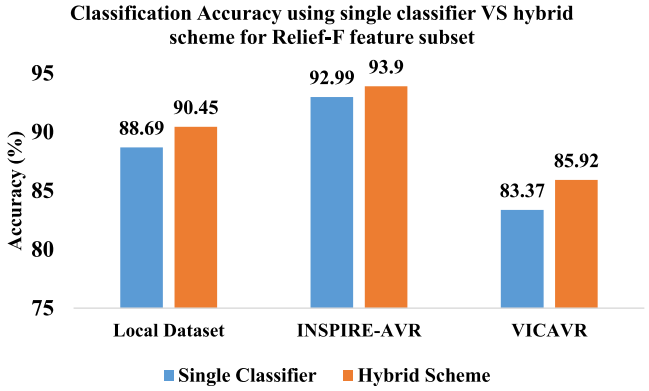


Fig. 20. Comparison of classification accuracy obtained using single classifier (shown in blue) and proposed hybrid classification (shown in orange), for feature subset from Relief-F ranking.

the segmented vessel network (Fig. 22(b)), candidate vessel segments for retinal vessel classification inside RoI (Fig. 22(c)) and classified vessel segments illustrated in Fig. 22(d), (blue and red circles show vessels classified correctly as vein and artery, respectively, whereas white and green circles show unclassified vessel segments for which ground truth is not available and vein segments misclassified as artery, respectively.

4.5. *Comparison of vessel classification accuracy with other state-of-art approaches*

Table 7 tabulates the techniques and outputs reported by previous researchers for retinal vessel classification using INSPIRE-AVR and VICAVR datasets and makes a comparison with the results shown by our system. We used accuracy, sensitivity and specificity metrics for comparison with vessel recognition results presented by already existing alternative methods. For comparison purpose with previous approaches, we selected the vessel classification results using hybrid labeling scheme with feature selection method that shows highest-performance. Highest-vessel classification accuracy is achieved using Relief-F ranking method with hybrid classification approach for local dataset (90.45%) and INSPIRE-AVR dataset (93.90%), while for VICAVR database, the highest accuracy (87.82%) is obtained using PCC method with hybrid classification method. For comparison of VICAVR results, among the opinion of three experts, we used the ground truth provided by Expert 1 with the dataset.[30]

Area Under Curve (AUC) metric is used by Niemeijer *et al.*,[33] for evaluation of their proposed vessel classification approach, however, in order to facilitate the comparison with vessel classification results of our system, we use the sensitivity and specificity values approximated by Dashtbozorg *et al.*[38] from the AUC reported by Niemeijer *et al.*[33] Note from Table 7 that both the sensitivity and specificity values presented by our system show a considerable superiority to the ones approximated for.[33] It should be noted that[33] vessel classification results are only for vessel center-line pixels whereas our method is evaluated on complete vessel segments in RoI. On the other hand, our method has not considered the small capillaries in vessel classification task due to the unavailability of ground truth. Although, the vessel center-line features used in Ref. 33 are also included in this research, but our method exploits comparatively a larger number of color planes for vessel characterization from both healthy and pathological-diseased images. Note that this comparison is with the approximated sensitivity and specificity values for vessel classification presented by Niemeijer *et al.*[33]

In comparison with vessel classification results proposed by,[38] they acquired 91.1% accuracy for INSPIRE-AVR dataset using 2-fold cross-validation with Linear Discriminant Analysis (LDA) classifier while our method achieved an accuracy of 92.26% using 10-fold cross-validation with fusion of classifier decisions. In addition, note that[38] achieved this accuracy using only a set of 19 features whereas our system obtained the mentioned accuracy using an optimal set of 37 features. In this case, although

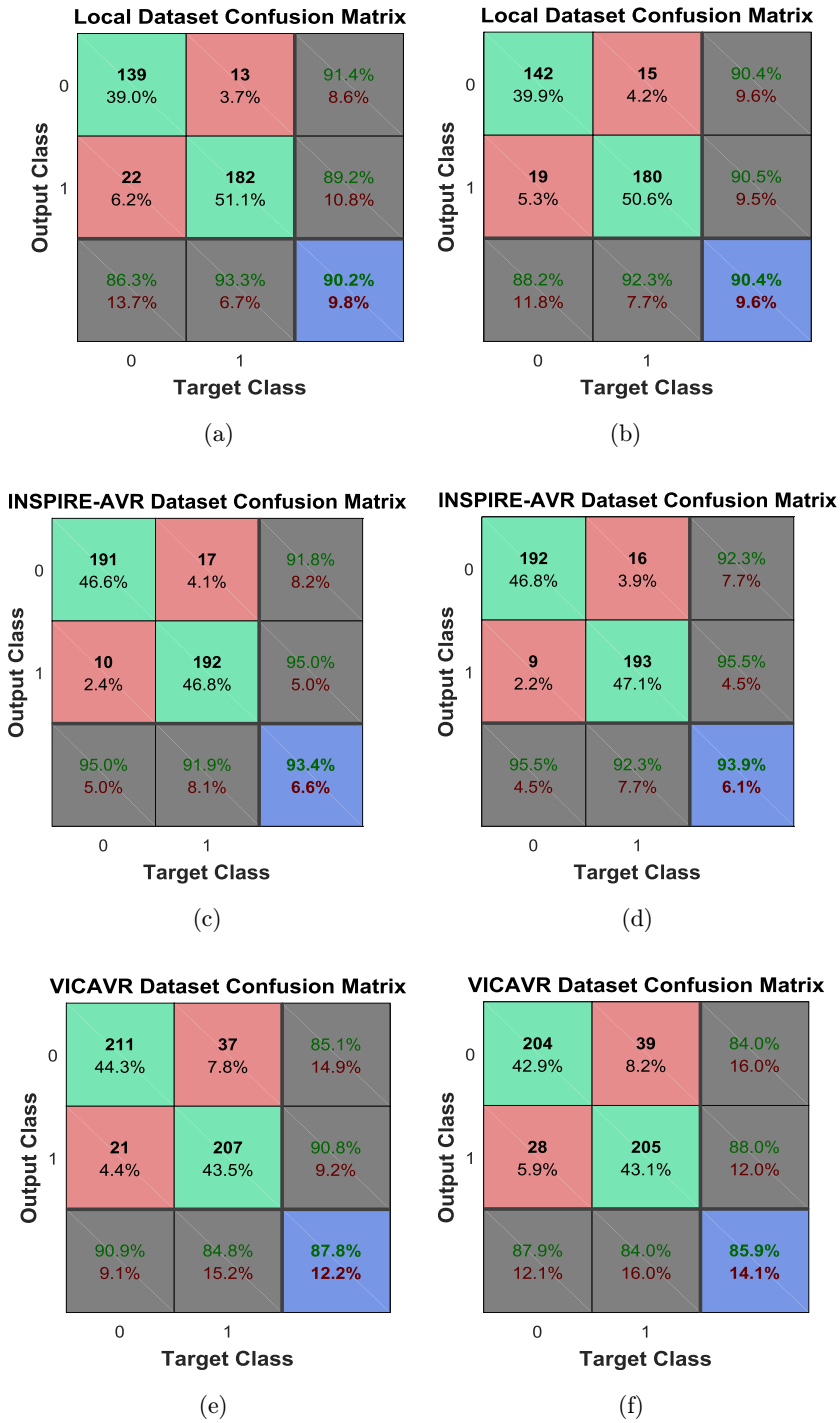
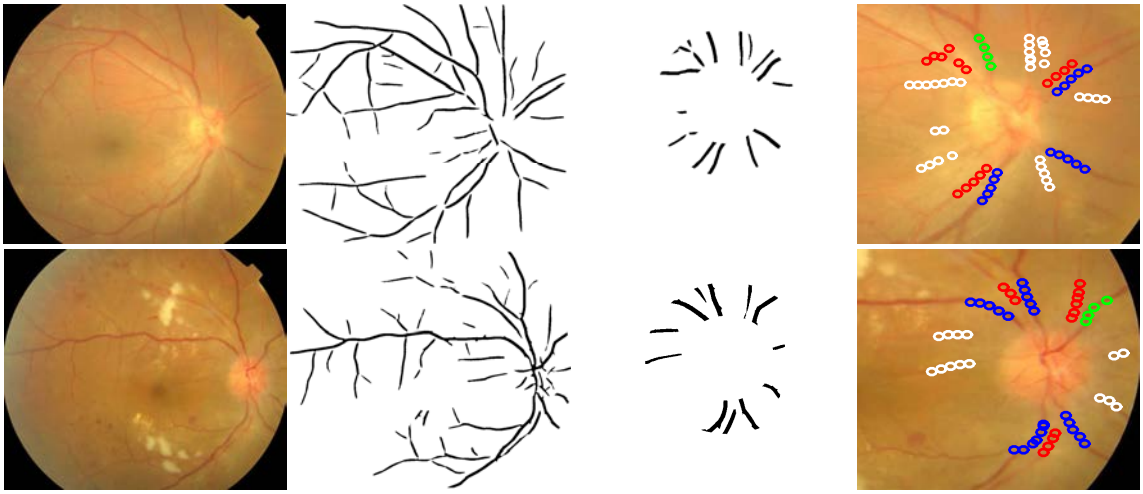


Fig. 21. Confusion matrices obtained for hybrid classification using features selected via PCC and Relief-F method. Left side: (a), (c) and (e) Vessel labeling performance with optimal feature subsets using PCC method. Right side: (b), (d) and (f) Vessel labeling performance with optimal feature subsets using PCC method.

our system achieved slightly higher vessel classification accuracy, sensitivity and specificity but if the number of features are compared, our system is inferior in efficiency as compared to the ones presented by Dashtbozorg *et al.*[38] The vessel classification approach presented by Dashtbozorg *et al.*[38] is tested on three databases and the results are reported for vessels inside RoI as well as complete retinal vessel tree. However, for computation of AVR, the classification of complete vessel tree is



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

Fig. 22. Classification of retinal vessels (images taken from local database). (a) Original retinal images, (b) Vessel segmentation results, (c) Vessels inside RoI and (d) Classified retinal vessel portions (blue: vein, red: artery, white: neither artery or vein (unclassified), green: vein that has been wrongly classified as artery).

not necessary. Classification of complete vessel tree is computationally expensive as compared to classification of small vessel portions inside RoI. In our method, only the vessels inside RoI are used, thus reducing the processing of redundant pixels that do not contribute. In order to have system with less complexity and increased efficiency, only the vessels inside RoI are assessed.

An accuracy of 87.6% is reported by Relan *et al.*[50] for the INSPIRE-AVR dataset as compared to 92.44% accuracy in our research, however, the sample size used in their paper is comparatively larger. Two features, i.e., mean of Red and Green channel from center-line pixels used by Relan *et al.*[50] for vessel recognition are also incorporated in our method. It is worth noting that none of these studies consider pathologically diseased images for evaluation of their proposed methods. For VICAVR, our system showed comparatively less classification accuracy. An accuracy of 88.8% is achieved by Vázquez *et al.*,[30] on VICAVR dataset which is a little higher than the ones achieved by our system. Recently, Vijayakumar *et al.*,[52] reported an accuracy of 92.4% on VICAVR dataset, however, they have not mentioned the number of vessels used for evaluation of their method. An apparent reason for low classification rate achieved using VICAVR can be the difference of ground-truth used for evaluation purpose.

The classification results presented for both healthy image database and diseased image database prove the capability of system in recognizing the vessels with higher accuracy.

4.6. *AVR computation results*

For evaluation purpose, the values of AVR calculated using the proposed method on three databases is compared with the corresponding AVR values in groundtruth. For INSPIRE-AVR database, we use the AVR values provided by both the observers 1 and 2 as benchmark for comparison and error calculation, since AVR estimated by an individual observer is dependent on his visual perception. We have used three parameters to assess the validity of our method in calculating AVR: (a) mean of difference between AVR values calculated by our system and those estimated by human expert, (b) mean of ratio of AVRs calculated automatically by our method to that of estimated manually, (c) mean of difference between AVRs by automatic method and average of (AVRs by observer 1 and 2). The parameter (a) is calculated by taking the mean of difference between AVRs generated by automatic system with the manual estimations. Parameter (b) is calculated by taking the ratio of AVRs calculated automatically to the corresponding AVRs given by human observer and then taking the mean of all those ratios. The closeness of parameter (b) which is basically mean of ratios with 1 depicts the closeness of AVRs automatically calculated and manually estimated. Parameter (c) is only calculated for INSPIRE-AVR

Table 7. Comparison of vessel classification with previous methods.

Method	Dataset	Technique	Performance parameter		
			Accuracy	Sensitivity	Specificity
Niemeijer *et al.*[33]	INSPIRE-AVR database (All center-line pixels detected in RoI)	27 features from Red, Green, Hue, Saturation and Intensity plane + prior information of retinal vessel arrangement + LDA classifier	—	∼78	∼78
Dashtbozorg *et al.*[38]	INSPIRE-AVR Database (All vessel segments inside RoI)	19 features based on HSI and RGB color channel + LDA classifier	91.1	91	86
	VICAVR dataset (All vessel segments including unclassified vessels)		89.8	—	—
Relan *et al.*[50]	INSPIRE-AVR dataset (483 vessel segments)	4 features from Green, Red and Hue planes + Squared-loss Mutual Information clustering	87.6	—	—
Vázquez *et al.*[30]	VICAVR database (All vessel segments including unclassified vessels)	5 features from vessel profiles in RGB, HSL and gray-level color space with clustering and tracking approach	88.8	—	—
V. Vijayakumar *et al.*[52]	VICAVR database	29 features from RGB, LAB, and YCbCr + SVM classifier	92.4	—	—
Our method	INSPIRE-AVR dataset (total 410 vessel segments inside RoI, unclassified vessel segments not included)	53 features including 18 features from gradient magnitude, 16 features from pixel raw values, 12 from histogram based and 6 from filter responses + Relief-F ranking followed by feature selection using three classifiers + decision fusion scheme	93.90	95.52	92.34
	Local Database (total 356 vessel segments inside RoI)	59 features including 20 features from gradient magnitude, 19 features from pixel raw values, 14 from histogram based and 5 from filter responses + Relief-F ranking followed by feature selection using three classifiers + decision fusion scheme	90.45	90.45	90.45
	VICAVR (total 476 vessel segments, unclassified vessel segments not included)	37 features including 4 features from gradient magnitude, 16 features from pixel raw values, 10 from histogram based and 6 from filter responses + PCC ranking followed by feature selection using three classifiers + decision fusion scheme	87.82	85.08	90.79

dataset. The use of these parameters has been motivated by Niemeijer *et al.*,[33] and Dashtbozorg, Mendonça, Campilho,[38] who evaluated their AVR calculation methods using parameter (a), while the parameter (b) has been used by Ruggeri *et al.*[88] in order to analyze the closeness of AVRs obtained automatically with those estimated manually. Parameter (c) is proposed here in order to analyze the difference in AVR values by proposed method with the average of AVRs by two experts. By taking the difference with average of AVRs estimated by two different observers, the bias towards the AVR values estimated by one person can be reduced. In order to carry out the comparative analysis, two parameters, i.e., (a) and (b), are also calculated for AVRs estimated by observer 1 and 2 for INSPIRE-AVR database to obtain inter-observer variability. This will allow to analyze the closeness of error between (automatic and manual AVRs) and (AVRs estimated by observer 1 and 2).

Table 8 shows the performance of our system in AVR calculation and comparison with manual estimations for INSPIRE-AVR database. Note that there is considerable inter-observer difference in AVRs estimated by two observers. The AVR values calculated by our method are relatively closer to the AVRs estimated by observer 1 as compared to the AVRs estimated by observer 2. It is also important to mention that our method achieved mean error of 0.0565, that is close to the inter-observer error, i.e., 0.0520. When the mean of difference between automatically calculated AVRs and average of AVRs is taken, the mean error declines to 0.0477.

Table 9 shows the results of AVR calculation for local and VICAVR databases, where the mean error is 0.0650 and 0.0849, respectively, larger than that calculated for INSPIRE-AVR database. Since the CAD systems are employed for acquiring diagnostic assessment, the presented results in Tables 8 and 9 demonstrate that the AVRs computed by our system can be considered as second independent opinion obtained from an automatic "Machine Expert". Furthermore, from the results of AVRs calculation on pathologically diseased database, it is revealed that our method is appreciable in providing the suitable AVR approximations.

4.7. *AVR calculation comparison with previous methods*

Table 10 shows the comparison of the AVRs calculated using our method with the results proposed by other authors. The AVR calculation performance in Ref. 33 was evaluated on INSPIRE-AVR dataset by comparing the AVRs resulted from their method with both observer 1 and 2. As shown in Table 10, the mean errors reported by Ref. 33 with respect to observer 1 and 2 are 0.06 and 0.05, respectively, whereas these values are 0.0565 and 0.074 using our method.

Table 8. AVR calculation results for INSPIRE-AVR database.

Index	Performance parameters	Values
(a)	Mean of difference between AVR Values by (automatic method and observer 1)	0.0565
	Mean of difference between AVR Values by (automatic method and observer 2)	0.0740
	Mean of difference in AVR Values by (observer 1 and 2)	0.0520
(b)	Mean of ratios of Automatic AVRs to Manually estimated AVRs by Observer 1	1.0030
	Mean of ratios of Automatic AVRs to Manual estimated AVRs by Observer	1.0129
	Mean of ratios of AVRs estimated (by observer 1 to AVRs by observer 2)	1.0099
(c)	Mean of difference of average of AVRs (estimated by observer 1 and 2) and AVRs by automatic method	0.0477

Table 9. AVR Results for local and VICAVR database.

Dataset	Performance parameters	Values
Local database	Mean error between AVR Values by (automatic method and expert)	0.0650
	Mean of ratios of Automatic AVRs to Manually estimated AVRs	1.0159
VICAVR database	Mean error between AVR Values by (automatic method and expert)	0.0849
	Mean of ratios of Automatic AVRs to Manually estimated AVRs	1.0623

Table 10. Comparison of AVR estimated by the proposed method with previously proposed results.

Method	Dataset	Performance parameter	Value
Niemeijer *et al.*[33]	INSPIRE-AVR	Mean Error between AVR values (computed automatically and provided by observer 1)[1]	0.06[1]
		Mean Error between AVR values (computed automatically and provided by observer 2)[2]	0.052[2]
Dashtbozorg *et al.*[38]	INSPIRE-AVR	Mean Error between AVR values (computed automatically and observer 1)[1]	0.05
Mendonça *et al.*[40]	INSPIRE-AVR	Mean Error between AVR values (computed automatically and observer 1)[1]	0.07
		Mean Error between AVR values (computed by semi-automated method and observer 1)[1]	0.04
Proposed System	INSPIRE-AVR dataset	Mean Error between AVR values (computed automatically and observer 1)[1]	0.0565[1]
		Mean Error between AVR values (computed automatically and observer 2)[2]	0.0740[2]
	Local Database	Mean Error between AVR values (computed automatically and estimated by human expert)	0.0650
	VICAVR dataset	Mean Error between AVR values (computed automatically and estimated by expert)	0.0849

Although our method showed slight improvement in reducing mean error between AVRs automatically calculated and estimated by observer 1 than the ones given in Ref. 33, even then our method is still appreciable. As shown in Table 10,[89] evaluated their method using INSPIRE-AVR dataset by considering AVR annotations only by observer 1. When compared with the results proposed by Dashtbozorg *et al.*[89] on INSPIRE-AVR database, our method showed approximately similar mean error. The third comparison is made with the method by Mendonça *et al.*,[40] in which automatic as well as semi-automated methods are used for AVR calculation and results are evaluated using the AVRs estimated by observer 1. Our method showed considerable superiority in calculating AVR values as compared to the AVRs calculated by Mendonça *et al.*,[40] using automatic method. The comparison results in Table 10 show that our method surpasses the results reported by Niemeijer *et al.*[33] and Mendonça *et al.*,[40] and presented similar mean error when compared to the AVR results reported by Dashtbozorg *et al.*[89]

5. Discussion and Conclusion

The proposed methodology includes seven modules: (a) Automatic detection and segmentation of retinal vessels; (b) Extraction of novel feature set to categorize vessels; (c) Ranking of features by Pearson Correlation Coefficient and Relief-F method; (d) Selection of features from ranked feature lists based on classification accuracy of three classifiers; (e) Classification of vessels by hybrid classification framework using selected feature subset; (f) Calculation of width of vessels and (g) Calculation of Arteriovenous Ratio. Particularly, two feature ranking techniques (PCC and Relief-F) followed by three classifiers for selection of features with multi-decision combination method for retinal vessel classification and subsequent AVR calculation, are evaluated in this paper for an improved HR detection system. The effect of specific feature ranking techniques with the use of multiple classifiers for feature selection and incorporation of "joint strength" of three supervised prediction models has not been evaluated in the past, therefore, the results obtained by the experiments can be used as a baseline or reference for future research.

The proposed methodology offers comparable results and works robustly on three databases acquired from different fundus cameras with different settings. The experimental evaluations highlight the strength of proposed vessel recognition model in capturing the relation between input features and classification outcomes effectively. Particularly, the arrangement of features and combinations of subsets according to feature lists ranked by PCC and Relief-F method have contributed to increase the retinal vessel classification accuracy, as compared to

performance of features without ranking. The AVR computed using the proposed method agrees with the manually estimated AVR with an error of 0.0650, 0.0565 and 0.0849 for local database, INSPIRE-AVR and VICAVR database, respectively. Agreement observed in the experiment results for INSPIRE-AVR database is comparable to the inter-observer variability where the difference between AVR estimations of two experts is 0.0520. This difference exists due to visual perception discrimination present in the observers. It can be seen from Table 9 that the algorithm detects the subjects suffering from HR with less error using images of low resolution and containing multiple retinal pathologies. It is successfully demonstrated that generalized arteriolar narrowing in retinal images can be quantified using the presented computer-aided process which may offer an opportunity for reduction of disease progression The experimental results show AVR as a significant indicator for prediction of HR in an individual. The system does not require any complex computation; however, one of the important limitations of the proposed algorithm is its dependency on the vessel segmentation results. Retinal vessels are known to be deteriorated in higher grades of HR which affects the vessel delineation process. Therefore, enhancing the performance of vessel segmentation is likely to improve the classification process, providing more efficient and robust computer-aided analysis system. Another constraint of our method is the limit in vessel classification performance due to presence of retinal pathologies since these pathologies may influence reflectivity. Although the presence of retinal pathologies is viewed to be less problematic in our research, the vessel classification and AVR measurement error remained high for images with the pathological structures. The segmentation of these pathologies from retinal images can contribute to the vessel classification effectiveness. Moreover, the sample size in our proposed research is comparatively small, especially for higher grades of HR. The experimental evaluations are promising; however, in this research we have not used all types of features for vessel classification and it requires more independent testing for validation of the fused classification scheme. Meanwhile, deep learning technology can be the way to achieve accurate vessel segmentation, which will form our future research. The structural information of vessels may allow acquiring better classification accuracy.

Acknowledgements

Samra Irshad has made the same contributions in writing this paper as the first author.

The author would like to thank Dr. Umer Salman from Hameed Latif Hospital, Lahore, Pakistan for assisting in providing ground truth for retinal vessel classification and AVR measurements in datasets.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. T. Y. Wong, R. McIntosh, "Systemic associations of retinal microvascular signs: A review of recent population-based studies", *Ophthalmic Physiol. Opt.* **25**(3), 195–204 (2005).
2. B. R. McClintic, J. I. McClintic, J. D. Bisognano, R. C. Block, "The relationship between retinal microvascular abnormalities and coronary heart disease: A review", *The Am. J. Med.* **123**(4), 374.e1–374.e7 (2010).
3. M. C. Knudtson, K. E. Lee, L. D. Hubbard, T. Y. Wong, R. Klein, B. E. K. Klein, "Revised formulas for summarizing retinal vessel diameters", *Current Eye Res.* **27**(3), 143–149 (2003).
4. M. K. Ikram, F. J. de Jong, J. R. Vingerling, J. C. Witteman, A. Hofman, M. M. Breteler, P. T. de Jong, "Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The Rotterdam Study", *Investigative Ophthalmol. Visual Sci.* **45**, 2129–2134 (2004).
5. J. T. DellaCroce, A. T. Vitale, "Hypertension and the eye", *Current Opinion in Ophthalmology*, **19**(6), 493–498 (2008).
6. M. O. M. Tso, L. M. Jampol, "Pathophysiology of hypertensive retinopathy", *Ophthalmology*, **89**(10), 1132–1145 (1982).
7. H. G. Scheie, "Evaluation of ophthalmoscopic changes of hypertension and arteriolar sclerosis", *AMA Archive. Ophthalmol.* **49**(2), 117–138 (1953).
8. N. M. Keith, H. P. Wagener, N. W. Barker, "Some different types of essential hypertension: Their course and prognosis", *The Am. J. Med. Sci.* **268**(6), 336 (1974).
9. T. Y. Wong, P. Mitchell, "Hypertensive retinopathy", *New England J. Med.* **351**(22), 2310–2317 (2004).
10. A. Grosso, F. Veglio, M. Porta, F. M. Grignolo, T. Y. Wong, "Hypertensive retinopathy revisited: Some answers, more questions", *British J. Ophthalmol.* **89**(12), 1646–1654 (2005).

11. N. L. Stokoe, R. W. Turner, "Normal retinal vascular pattern. Arteriovenous ratio as a measure of arterial calibre", *The British J Ophthalmol.* **50**(1), 21 (1966).
12. T. Y. Wong, M. D. Knudtson, R. Klein, B. E. Klein, S. M. Meuer, L. D. Hubbard, "Computer-assisted measurement of retinal vessel diameters in the Beaver Dam Eye Study: Methodology, correlation between eyes, and effect of refractive errors", *Ophthalmology*, **111**(6), 1183–1190 (2004).
13. N. Patton, T. M. Aslam, T. MacGillivray, I. J. Deary, B. Dhillon, R. H. Eikelboom, K. Yogesan, I. J. Constable, "Retinal image analysis: Concepts, applications and potential", *Progress in Retinal and Eye Res.* **25**(1), 99–127 (2006).
14. F. C. Delori, E. S. Gragoudas, R. Francisco, R. C. Pruett, "Monochromatic ophthalmoscopy and fundus photography: The normal fundus", *Archives of ophthalmol.* **95**(5), 861–868 (1977).
15. M. A. Pérez, B. B. Bruce, N. J. Newman, V. Biousse, "The use of retinal photography in non-ophthalmic settings and its potential for neurology", *The Neurologist* **18**(6), 350 (2012).
16. B. B. Bruce, C. Lamirel, D. W. Wright, A. Ward, K. L. Heilpern, V. Biousse, N. J. Newman, "Nonmydriatic ocular fundus photography in the emergency department", *New England J. of Med.* **364**(4), 387–389 (2011).
17. H. Li, W. Hsu, M. L. Lee, T. Y. Wong, "Automatic grading of retinal vessel caliber", *IEEE Trans. Biomed. Eng.* **52**(7), 1352–1355 (2005).
18. G. C. Manikis, V. Sakkalis, X. Zabulis, P. Karamaounas, A. Triantafyllou, S. Douma, C. Zamboulis, K. Marias, "An image analysis framework for the early assessment of hypertensive retinopathy signs", in *E-Health and Bioengineering Conf. EHB*, pp. 1–6, IEEE (2011).
19. K. Narasimhan, V. C. Neha, K. Vijayarekha, "Hypertensive retinopathy diagnosis from fundus images by estimation of AVR", *Procedia Eng.* **38**, 980–993 (2012).
20. S. Vázquez, N. Barreira, M. Penedo, M. Rodriguez-Blanco, F. Gómez-Ulla, A. González, G. Coll de Tuero, Automatic arteriovenous ratio computation: Emulating the experts, *Technological Innovation for Value Creation*, pp. 563–570 (2012).
21. X. Zabulis, A. Triantafyllou, P. Karamaounas, C. Zamboulis, S. Douma, An image analysis system for the assessment of retinal microcirculation in hypertension and its clinical evaluation, in *XIII Mediterranean Conf. Medical and Biological Engineering and Computing* pp. 330–335, Springer (2014).
22. V. Joshi, C. Agurto, E. S. Barriga, S. Nemeth, P. Soliz, Clinical utilization of automated image analysis software for improving retinal reader's performance. in *Southwest Symp. Image Analysis and Interpretation SSIAI*, pp. 145–148, IEEE (2016).
23. X. Yin, B. W. Ng, J. He, Y. Zhang, D. Abbott, "Accurate image analysis of the retina using hessian matrix and binarisation of thresholded entropy with application of texture mapping", *PloS one*, **9**(4), e95943 (2014).
24. Y. Zhao, L. Rada, K. Chen, S. P. Harding, Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images", *IEEE Trans. Med. Imag.* **34**(9), 1797–1807 (2015).
25. D. Pandey, X. Yin, H. Wang, Y. Zhang, "Accurate vessel segmentation using maximum entropy incorporating line detection and phase-preserving denoising", *Comput. Vision and Image Understand.* **155**, 162–172 (2017).
26. A. M. Aibinu, M. I. Iqbal, A. A. Shafie, M. J. E. Salami, M. Nilsson, "Vascular intersection detection in retina fundus images using a new hybrid approach", *Comput. Biol. Med.*, **40**(1), 81–89 (2016).
27. G. Azzopardi, N. Petkov, "Automatic detection of vascular bifurcations in segmented retinal images using trainable COSFIRE filters", *Pattern Recogn. Lett.* **34**(8), 922–933 (2013).
28. D. Ortíz, M. Cubides, A. Suárez, M. Zequera, J. Quiroga, J. Gómez, N. Arroyo, Support system for the preventive diagnosis of hypertensive retinopathy, in *37th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 5649–5652. IEEE (2010).
29. R. A. Welikala, M. M. Fraz, S. Hayat, A. R. Rudnicka, P. J. Foster, P. H. Whincup, C. G. Owen, D. P. Strachan, S. A. Barman, Automated retinal vessel recognition and measurements on large datasets, in *37th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 5239–5242, IEEE (2015).
30. S. G. Vázquez, B. Cancela, N. Barreira, M. G. Penedo, M. Saez, On the automatic computation of the arterio-venous ratio in retinal images: Using minimal paths for the artery/vein classification, in *Int. Conf. Digital Image Computing: Techniques and Applications DICTA*, pp. 599–604. IEEE (2010).
31. S. G. Vázquez, B. Cancela, N. Barreira, M. G. Penedo, M. Rodríguez-Blanco, M. P. Seijo, G. C. de Tuero, M. A. Barceló and M. Saez, "Improving retinal artery and vein classification by means of a minimal path approach", *Machine vision and applications*, **24**(5), 919–930 (2013).
32. Q. Mirsharif, F. Tajeripour, H. Pourreza, "Automated characterization of blood vessels as

arteries and veins in retinal images", *Comput. Med. Imag. Graphic.* **37**(7), 607–617 (2013).
33. M. Niemeijer, X. Xu, A. V. Dumitrescu, P. Gupta, B. van Ginneken, J. C. Folk, M. D. Abramoff, "Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs", *IEEE Trans. Med. Imag.* **30**(11), 1941–1950 (2011).
34. K. Rothaus, X. Jiang, P. Rhiem, "Separation of the retinal vascular graph in arteries and veins based upon structural knowledge", *Image Vision Comput.* **27**(7), 864–875 (2009).
35. Y. Huang, J. Zhang, Y. Huang, "An automated computational framework for retinal vascular network labeling and branching order analysis", *Microvascular Res.* **84**(2), 169–177 (2012).
36. M. Saez, S. González-Vázquez, M. González-Penedo, M. A. Barceló, M. Pena-Seijo, G. C. de Tuero, A. Pose-Reino, "Development of an automated system to classify retinal vessels into arteries and veins", *Comput. Methods Programs Biomed.* **108**(1), 367–376 (2012).
37. R. Estrada, M. J. Allingham, P. S. Mettu, S. W. Cousins, C. Tomasi, S. Farsiu, "Retinal artery-vein classification via topology estimation", *IEEE Trans. Medical Imag.* **34**(12), 2518–2534 (2015).
38. B. Dashtbozorg, A. M. Mendonça, A. Campilho, "An automatic graph-based approach for artery/vein classification in retinal images", *IEEE Trans. Image Process.* **23**(3), 1073–1083 (2014).
39. B. Dashtbozorg, A. M. Mendonça, A. Campilho, Automatic classification of retinal vessels using structural and intensity information, in *Iberian Conf. Pattern Recognition and Image Analysis*, pp. 600–607, Springer (2013).
40. A. M. Mendonça, B. Remeseirob, B. Dashtbozorgc, A. Campilhoa, Automatic and semi-automatic approaches for arteriolar-to-venular computation in retinal photographs, in *Proc. of SPIE Vol*, Vol. 10134, pp. 101341L-1 (2017).
41. X. Xu, W. Ding, M. D. Abràmoff, R. Cao, "An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image", *Comput. Meth. Program. Biomed.* **141**, 3–9 (2017).
42. G. Litjens *et al.*, "A survey on deep learning in medical image analysis", *Med. Image Anal.* **42**, 66–68 (2017).
43. Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images", *IEEE Trans. Med. Imag.* **35**(1), 109–118 (2016).
44. Z. Liu, Q. Pan, J. Dezert, J.-W. Han, Y. He, "Classifier fusion with contextual reliability evaluation", *IEEE Trans. Cybernetic.* **48**(5) 1605–1618 (2018).
45. M. Graña, M. Termenon, A. Savio, A. Gonzalez-Pinto, J. Echeveste, J. M. Pérez, A. Besga, "Computer aided diagnosis system for Alzheimer disease using brain diffusion tensor imaging features selected by Pearson's correlation", *Neurosci. Lett.* **502**(3), 225–229 (2011).
46. N. P. Pérez, M. A. G. López, A. Silva, I. Ramos, "Improving the Mann–Whitney statistical test for feature selection: An approach in breast cancer diagnosis on mammography", *Artific. Intell. Med.* **63**(1), 19–31 (2015).
47. I. Beheshti, H. Demirel, H. Matsuda, "Alzheimer's disease neuroimaging initiative. Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm," *Comput. Biol. Med.* **83** 109–119 (2017).
48. D. Relan, T. MacGillivray, L. Ballerini, E. Trucco, Retinal vessel classification: Sorting arteries and veins, in *35th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 7396–7399, IEEE (2013).
49. D. Relan, T. MacGillivray, L. Ballerini, E. Trucco, Automatic retinal vessel classification using a least square-support vector machine in VAMPIRE, in *36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 142–145. IEEE (2014).
50. D. Relan, L. Ballerini, E. Trucco, T. MacGillivray, Retinal vessel classification based on maximization of squared-loss mutual information, in *Machine Intelligence and Signal Processing*, pp. 77–84, Springer (2016).
51. C. Agurto, V. Joshi, S. Nemeth, P. Soliz, S. Barriga, Detection of hypertensive retinopathy using vessel measurements and textural features, in *36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 5406–5409, IEEE (2014).
52. V. Vijayakumar, D. D. Koozekanani, R. White, J. Kohler, S. Roychowdhury, K. K. Parhi, Artery/vein classification of retinal blood vessels using feature selection, in *38th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 1320–1323, IEEE (2016).
53. M. M. Fraz, A. R. Rudnicka, C. G. Owen, D. P. Strachan, S. A. Barman, Automated arteriole and venule recognition in retinal images using ensemble classification, in *Int. Conf. Computer Vision Theory and Applications VISAPP*, pp. 194–202. IEEE (2014).

54. G. Niu, T. Han, B. S. Yang, A. C. C. Tan, "Multi-agent decision fusion for motor fault diagnosis", *Mechanic. Syst. Signal Process.* **21**(3), 1285–1299 (2007).
55. P. Du, W. Zhang, H. Sun, "Multiple classifier combination for hyperspectral remote sensing image classification", *Multiple Classifier Syst.*, 52–61 (2009).
56. M. Liu, D. Zhang, D. Shen, "Hierarchical fusion of features and classifier decisions for Alzheimer's disease diagnosis", *Human Brain Mapping* **35**(4), 1305–1319 (2014).
57. S. S. Devi, A. Roy, J. Singha, S. A. Sheikh, R. H. Laskar, "Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear", *Multimedia Tools and Appl.* 1–30 (2016).
58. M. G. Kendall, J. D. Gibbons, *Rank Correlation Methods*, 5th Edition. Oxford University Press, New York (1990).
59. I. Kononenko, Estimating attributes: Analysis and extensions of RELIEF, in *European Conf. Machine Learning*, pp. 171–182, Springer (1994).
60. I. Jamal, M. U. Akram, A. Tariq, "Retinal image preprocessing: Background and noise segmentation", *TELKOMNIKA (Telecommun. Comput. Electron. Control)*, **10**(3), 537–544 (2012).
61. T. S. Lee, "Image representation using 2D Gabor wavelets", *IEEE Trans. Pattern Anal. Machine Intell.* **18**(10), 959–971 (1996).
62. A. Usman, S. A. Khitran, M. U. Akram, Y. Nadeem, A robust algorithm for optic disc segmentation from colored fundus images, in *Int. Conf. Image Analysis and Recognition ICIAR*, pp. 303–310. Springer (2014).
63. S. Irshad, X. Yin, L. Q. Li, U. Salman, Automatic optic disk segmentation in presence of disk blurring, in *Int. Symp. Visual Computing*, pp. 13–23, Springer (2016b).
64. J. C. Parr, G. F. S. Spears, "Mathematic relationships between the width of a retinal artery and the widths of its branches", *Am. J. Ophthalmol.* **77**(4), 478–483 (1974).
65. L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, L. S. Cooper, A. R. Sharrett, M. D. Davis, J. Cai, "Atherosclerosis risk in communities study group, methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study", *Ophthalmology* **106**(12), 2269–2280 (1999).
66. M. Rumpf, A. Telea, A continuous skeletonization method based on level sets. *IEEE TCVG Symposium on Visualization*, pp. 151–158 (2002).
67. S. Irshad, M. U. Akram, S. Ayub, A. Ayaz, Retinal blood vessels differentiation for calculation of arterio-venous ratio, in *Int. Conf. Image Analysis and Recognition*, pp. 411–418, Springer (2015).
68. S. Irshad, M. Ahmad, M. U. Akram, A. W. Malik, S. Abbas, Classification of vessels as arteries verses veins using hybrid features for diagnosis of hypertensive retinopathy, in *Int. Conf. Imaging Systems and Techniques IST*, pp. 472–475, IEEE (2016).
69. T. Leung, J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons", *Int. J. Comput. Vision*, **43**(1), 29–44 (2001).
70. C. Schmid, Constructing models for content-based image retrieval, in *Computer Vision and Pattern Recognition CVPR*, IEEE (2001).
71. J. M. Geusebroek, A. W. Smeulders, J. Van De Weijer, "Fast anisotropic gauss filtering", *IEEE Trans. Image Processing* **12**(8), 938–943 (2003).
72. M. Douze, A. Ramisa, C. Schmid, Combining attributes and fisher vectors for efficient image retrieval, in *Computer Vision and Pattern Recognition CVPR*, pp. 745–752, IEEE (2011).
73. I. Dimitrovski, D. Kocev, I. Kitanovski, S. Loskovska, S. Džeroski, "Improved medical image modality classification using a combination of visual and textual features", *Comput. Med. Imaging Graphic.* **39**, 14–26 (2015).
74. H. Zhang, W. Gao, X. Chen, D. Zhao, "Object detection using spatial histogram features", *Image Vision Comput.* **24**(4), 327–341 (2006).
75. H. Breu, J. Gil, D. Kirkpatrick, M. Werman, "Linear time Euclidean distance transform algorithms," *IEEE Trans. Pattern Anal. Machine Intell.* **17**(5), 529–533 (1995).
76. J. Tang, S. Alelyani, H. Liu, Feature selection for classification: A review, *Data Classification: Algorithms and Applications*, pp. 37 (2014).
77. Y. Saeys, T. Abeel, Y. Van de Peer, "Robust feature selection using ensemble feature selection techniques", *Machine Learn. Knowledge Discovery in Databases*, 313–325 (2008).
78. I. Slavkov, B. Ženko, S. Džeroski, Evaluation method for feature rankings and their aggregations for biomarker discovery. in *Machine Learning in Systems Biology*, pp. 122–135 (2009).
79. I. Beheshti, H. Demirel, F. Farokhian, C. Yang, H. Matsuda, "Alzheimer's disease neuroimaging initiative, Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error," *Comput. Methods Programs Biomed.* **137**, 177–193 (2016).
80. S. Fakhraei, H. Soltanian-Zadeh, F. Fotouhi, "Bias and stability of single variable classifiers for feature

ranking and selection", *Expert Syst. Appl.* **41**(15), 6945–6958 (2014).
81. T. Cover, P. Hart, "Nearest neighbor pattern classification", *IEEE trans. Inf. theo.* **13**(1), 21–27 (1967).
82. C. Cortes, V. Vapnik, "Support-vector networks," *Machine Learn.* **20**(3), 273–297 (1995).
83. M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, C. F. Tsai, "SVM and SVM ensembles in breast cancer prediction", *PloS one* **12**(1), e0161501 (2017).
84. S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, Englewood Cliffs (1995).
85. D. Ruta, B. Gabrys, "Classifier selection for majority voting" *Information Fusion* **6**(1), 63–81 (2005).
86. K. Woods, W. P. Kegelmeyer, K. Bowyer, "Combination of multiple classifiers using local accuracy estimates", *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(4), 405–410 (1997).
87. J. Kittler, M. Hatef, R. P. Duin, J. Matas, "On combining classifiers", *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(3), 226–239 (1998).
88. E. Grisan and A. Ruggeri, A divide et impera strategy for automatic classification of retinal vessels into arteries and veins, *25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Cancun, Mexico, Vol 1 (2003).
89. B. Dashtbozorg, A. M. Mendonça, A. Campilho, An automatic method for the estimation of arteriolar-to-venular ratio in retinal images, in *2013 IEEE 26th Int. Symp. Computer-Based Medical Systems (CBMS),* pp. 512–513, IEEE (2013).