Pattern Recognition 46 (2013) 107–116

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

Identification and classification of microaneurysms for early detection of diabetic retinopathy

M. Usman Akram [a,*], Shehzad Khalid [b], Shoab A. Khan [a]

[a] Department of Computer Engineering, College of Electrical & Mechanical Engineering, National University of Sciences & Technology, Islamabad, Pakistan
[b] Department of Computer & Software Engineering, Bahria University, Islamabad, Pakistan

ARTICLE INFO

Article history:
Received 28 January 2012
Received in revised form
1 July 2012
Accepted 7 July 2012
Available online 17 July 2012

Keywords:
Medical image processing
Diabetic retinopathy
Microaneurysms
Classification
m-Mediods

ABSTRACT

Diabetic retinopathy is a progressive eye disease which may cause blindness if not detected and treated in time. The early detection and diagnosis of diabetic retinopathy is important to protect the patient's vision. The accurate detection of microaneurysms (MAs) is a critical step for early detection of diabetic retinopathy because they appear as the first sign of disease. In this paper, we propose a three-stage system for early detection of MAs using filter banks. In the first stage, the system extracts all possible candidate regions for MAs present in retinal image. In order to classify a candidate region as MA or non-MA, the system formulates a feature vector for each region depending upon certain properties, i.e. shape, color, intensity and statistics. We present a hybrid classifier which combines the Gaussian mixture model (GMM), support vector machine (SVM) and an extension of multimodel mediod based modeling approach in an ensemble to improve the accuracy of classification. The proposed system is evaluated using publicly available retinal image databases and achieved higher accuracy which is better than previously published methods.

1. Introduction

Diabetic retinopathy (DR) is nominated as the most common cause of blindness for the past 50 years. The epidemiological studies carried out in industrialized countries classify DR amongst the four main causes of sight problems over the whole population [1,2]. DR is an eye disease in which diabetes affects the blood vessels present in human retina which becomes one of the main sources of vision impairment [1]. One out of five patients with newly discovered type II diabetes has DR at the time of diagnosis whereas in the first five years after diagnosis of type I diabetes, DR almost never occurs. However after 15 years, almost all patients with type I and two-third of those with type II diabetes have signs of DR [1]. The common symptoms of diabetic retinopathy are blurred vision, floaters and flashes, and sudden loss of vision [3].

Early detection and treatment of DR is very crucial as it is a progressive disease and its severity is determined by the number and the types of lesions present in the fundus image. A healthy retina contains blood vessels, optic disc and macula as its main components and any changes in these components are signs of any eye disease [3]. DR is broadly divided into two stages, i.e. non-proliferative DR (NPDR) and proliferative DR (PDR). NPDR, also known as background DR, occurs when diabetes damages the blood vessels inside the retina causing leakage of blood and fluid on the surface of retina [4]. The functionality of retina reduces as it becomes wet and swollen due to this leakage. NPDR may contain different signs of retinopathy such as microaneurysms (MAs), hemorrhages (H), hard exudates (HE) and soft exudates or cotton wool spots (CWS) [5]. Depending upon the presence and quantity of these lesions, NPDR is further divided into three stages, i.e. mild, moderate and severe [5].

MAs are the first sign of NPDR and are caused by the focal dilatations of thin blood vessels. MAs are of small size, almost round in shape and red in color. The next sign of DR is H which is also referred to as dot or blot H. When the wall of thin vessels or MAs is sufficiently weakened, it may rupture and give rise to an H. Dot hemorrhages appear as bright small red dots and blot hemorrhages are larger red lesions. Sometimes dot hemorrhages and MAs are considered as a single red lesion class known as HMAs [2]. Fig. 1 shows retinal image with regions containing HMAs enlarged for ease of visualization.

Diabetes is very common disease and patient needs regular screening for early detection of DR. The ophthalmologists have to examine a large number of images in order to perform mass screening. The cost of manual examination is prohibiting the implementation of screening on a large scale. A possible solution could be the development of an automated screening system for

* Corresponding author: Tel.: +92 3336913921; fax: +92 512305071.
E-mail addresses: usmakram@gmail.com (M.U. Akram), shehzad_khalid@hotmail.com (S. Khalid), shoab@carepvtltd.com (S.A. Khan).

0031-3203/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2012.07.002

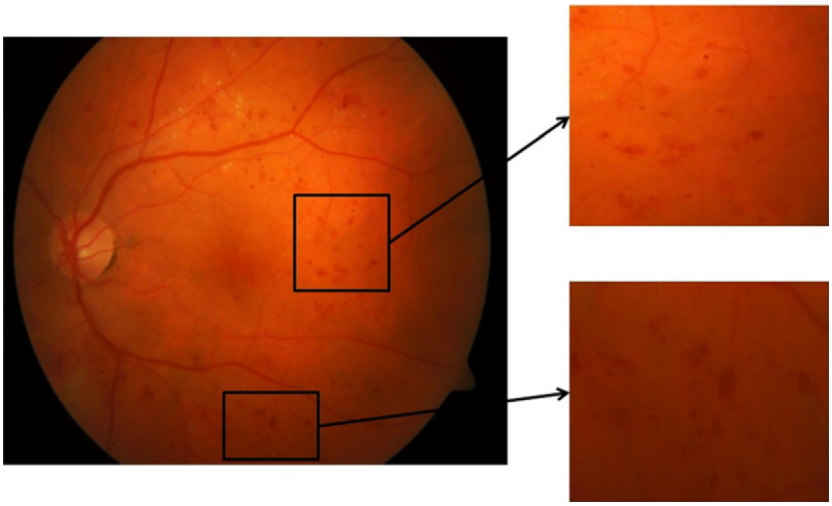

Fig. 1. Retinal image and enlarged regions containing HMAs.

retinal images [6]. Computer aided diagnostic systems for eye diseases use digital fundus images which are an essential mean to document and diagnose various eye diseases in clinics. Such a system should be able to distinguish between affected retinal images and normal retinal images. This will significantly reduce the workload for the ophthalmologists as they have to examine only those images diagnosed by the system as potential anomalies containing affected retina [7].

The remaining paper is organized as follows: Section 2 contains brief review of some recently published work relevant to DR. A flow diagram and overview of the proposed system is given in Section 3. It briefly discusses the complete systems and its different phases. Then in Sections 4 and 5, we present the detailed proposed techniques for candidate region detection and descriptions of all features which we compute for each candidate region. The modeling of proposed hybrid classifier is described in Section 6 followed by the experimental results and comparisons with previous techniques in Section 7. The last section summarizes the paper.

2. Related work

NPDR contains the early signs of DR and it is very crucial to detect and diagnose DR at its early stages. If a person with diabetes receives proper eye care regularly, and treatment when necessary, then DR will rarely cause total blindness. In the study of DR, most of the work is done to find MAs as they are early signs of the presence of retinopathy in fundus image. MAs appear as small circular dark spots on the surface of the retina. The detection of microaneurysms is still an open issue. Thus, several recent works focus on this problem, including an online challenge for MA detectors [3].

Recently, a number of studies being carried out for accurate detection of MAs, i.e. a successive clutter rejection based approach are presented by Ram et al. [8]. They designed a specialized feature based system that rejects specific classes of clutter while passing majority of true MAs. The potential MAs that remain after the final rejection stage are assigned a score which is based on their similarity to true MAs. Antal et al. [9] presented a method based on combinations of several preprocessing and candidate extractors. They used an ensemble learning based criteria for classification. A multi-scale correlation coefficients based method was proposed by Zhang et al. [10]. They used dynamic thresholding for accurate detection of MAs. A morphological operations based technique for detecting candidate MAs was proposed by Walter et al. [11]. They used just 20 images with total 297 manually detected MAs. A supervised density-based classifier was used for the MA classification. Quellec et al. [12] presented a template matching based method using an optimal wavelet transform. The matching was done using a generalized Gaussian template in wavelet domain. They evaluated their technique on 120 images with possible MAs. Some studies had used blood vessel segmentation for MAs detection. A local contrast normalization based method was proposed in [13]. They used watershed transform to detect MAs by distinguishing between MAs and other dots present on retina. The reported sensitivity and specificity were 85.4% and 83.1%, respectively. An online competition for MAs detection with the name of retinopathy online challenge (ROC) is introduced by the University of Iowa and ROC organizers [14]. The purpose is to improve the quality of computer aided and automated diagnosis of DR. The results of first international competition were reported in [15].

In studies of automated diagnosis for DR, hemorrhage and microaneurysm (HMA) are considered as a common class. A moat operator was used for automated detection of HMA and HE [16]. In this study only 30 retinal images were used for HMA and HE classification. A similar study was done by Usher et al. [17] by using retinal images form 1273 patients. Hatanaka et al. [18] did a study purely on H detection using 125 digital retinal images. They used hue, saturation and value (HSV) color space and Mahalanobis distances for classification of Acharya et al. [19] presented a higher order spectra and SVM based system for NPDR and PDR detection. They classified all the three stages of NPDR with accuracies of 90%, 85% and 70% for mild, moderate and severe NPDR, respectively. A computer aided system for classification of MAs, H, HE and CWS was proposed in [20]. They were able to identify HMA, HE and CWS with an accuracy of 82.6%, 82.6%, and 88.3%, respectively.

The automated systems for detection of MAs described in this section normally use different features and accurate classifier to improve the accuracy but still there is a space for improvement due to the existence of blood vessels. In our proposed system, we extract the vascular pattern first in order to improve the accuracy as MAs and blood vessels have almost similar intensity values which can lead to false MA regions. So in a combination of reliable feature set and classifier, we also keep in mind the fact that MAs have very close relation with blood vessels and the removal of vessel pixels has further improve the accuracy of MAs detection.

This paper describes a system for detection of MAs by extracting candidate regions. It generates a features set for each region depending upon their shape, gray level, color and statistical based properties. The true MA regions are selected and classified using a hybrid classifier which is a weighted combination of multivariate m-Mediods, GMM and SVM. The main contribution of the paper is that it improves the accuracies of all three stages using accurate vascular segmentation, sound feature set and hybrid classifier for accurate detection of MAs. The proposed method introduces the elimination of blood vessel pixels prior to MAs detection and enhances the feature space before applying the classifier. We also present the extension of m-Mediods model in hybrid classifier to improve the overall accuracy.

3. System overview

The diagnosis of DR is broadly divided into two categories, i.e. screening of DR and monitoring of DR. The automated system for screening of DR uses a fundus image as an input and find all lesions present in that image, whereas in monitoring, the system detects the changes between the two images of the same eye taken in different time moments [21]. This paper presents a method for the screening of DR and it identifies the first sign of DR that is MA using digital retinal images. Fig. 2 shows a complete flow diagram of the proposed system.

Overall, we divide the proposed system into three classical stages, i.e. candidate region extraction, feature vector formation and classification. The candidate region extraction phase enhances the regions containing red lesions and extracts all possible candidate regions for MAs. It also removes all possible blood vessel pixels to minimize spurious regions. In stage 2, feature vectors are calculated for each extracted region using some distinct properties. Finally, the classifier separates out the MA and non-MA regions.

4. Candidate region extraction

A reliable computer aided system should detect all potential regions in input fundus image for identification of MAs. The proposed algorithm for the candidate region extraction works in three phases. In phase 1, it improves the contrast of dark regions by using mathematical morphological operations, contrast normalization and filter banks. Phase 2 performs blood vessel enhancement and segmentation. The last phase of candidate region extraction eliminates all blood vessel pixels from candidate pixels in order to reduce the spurious regions and to increase the accuracy of classifier.

MAs appear as dark red dots and patches in fundus image with highest contrast in the green plane of the color image, respectively. Ophthalmologists have no difficulty in identifying these lesions but there are different factors such as variability of image clarity, variation in image background texture, confusion with blood vessel pixels and the presence of other lesions which make identification difficult for automated systems.

In the first phase, morphological opening is used to smooth OD and other bright lesions if present using the following equation [23]:

$$\phi_f^{(sB)} = \max[\min f(x+b)] \tag{1}$$

here f is the preprocessed colored image and $b \in sB$, where sB represents the structuring element (SE) B with size s. This gives us a smooth fundus region ϕ_f containing dark lesions and vessels only but they need contrast enhancement. The objective of contrast enhancement is to improve the contrast of lesions for easy detection using a $w \times w$ sliding window with assumption that w is large enough to contain a statistically representative distribution of the local variation of lesions [24].

$$g = 255\frac{[\Phi_w(\phi_f) - \Phi_w(\phi_{fmin})]}{[\Phi_w(\phi_{fmax}) - \Phi_w(\phi_{fmin})]} \tag{2}$$

where Φ_w is the sigmoid function for a window defined as

$$\Phi_w(\phi_f) = \left[1 + \exp\left(\frac{m_w - f}{\sigma_w}\right)\right]^{-1} \tag{3}$$

ϕ_{fmax} and ϕ_{fmin} are the maximum and minimum intensity values of smooth green channel image, respectively. m_w and σ_w are the mean and variance of intensity values within the window.

Fig. 3 shows the results of opening and contrast enhancement. g represents the contrast enhanced image which is given as an input to Gabor filter banks for enhancement of lesions. Gabor filters are famous due to their fine frequency tuning and orientation selectiveness. They are appropriate for texture representation and discrimination [25]. Gabor filter is represented by a Gaussian kernel function which can model a wide range of shapes depending upon the values of its parameters [25]. This property makes them suitable for MAs and dot hemorrhage detection

$$G(x,y,\sigma,\Omega,\theta,r) = \frac{1}{\sqrt{\pi r}\sigma} e^{-(1/2)[(d_1/\sigma)^2 + (d_2/\sigma)^2]}(d_1(\cos\Omega + \imath \sin\Omega)) \tag{4}$$

where σ, Ω and r are the standard deviations of Gaussian, spatial frequency and aspect ratio, respectively, θ is the orientation of filter and $d_1 = x\cos\theta + y\sin\theta$ and $d_2 = -x\sin\theta + y\cos\theta$. The contrast enhanced image g is convolved with Gabor filter G centered at location(s,t) to generate Gabor filter response γ for selected values of σ, Ω and θ is given in the following equation [25].

$$\gamma(\sigma,\Omega,\theta) = \sum_x \sum_y g(x,y)G(s-x,t-y,\sigma,\Omega,\theta,r) \tag{5}$$

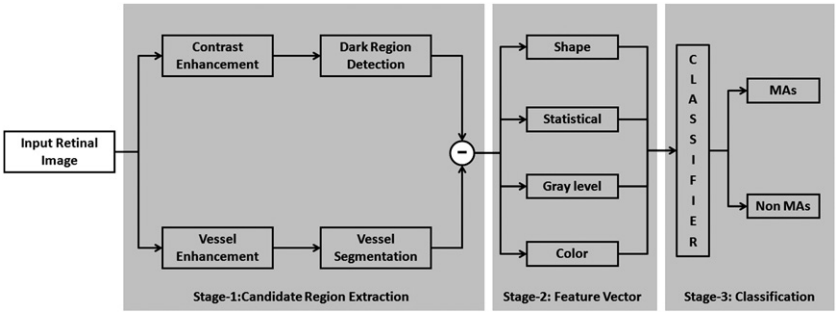

Fig. 2. Flow diagram for the proposed system.

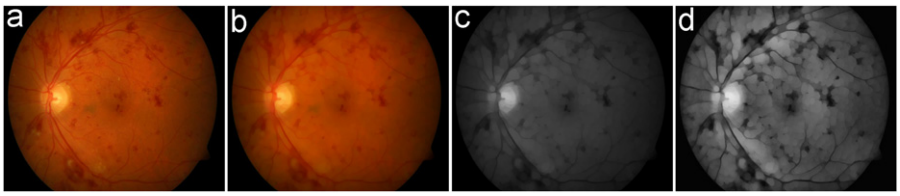

Fig. 3. Red lesion enhancement: (a) original retinal image; (b) smoothing of bright lesions and OD using opening morphological operation; (c) green channel obtained from (b); (d) contrast enhanced green channel highlighting MAs and H.

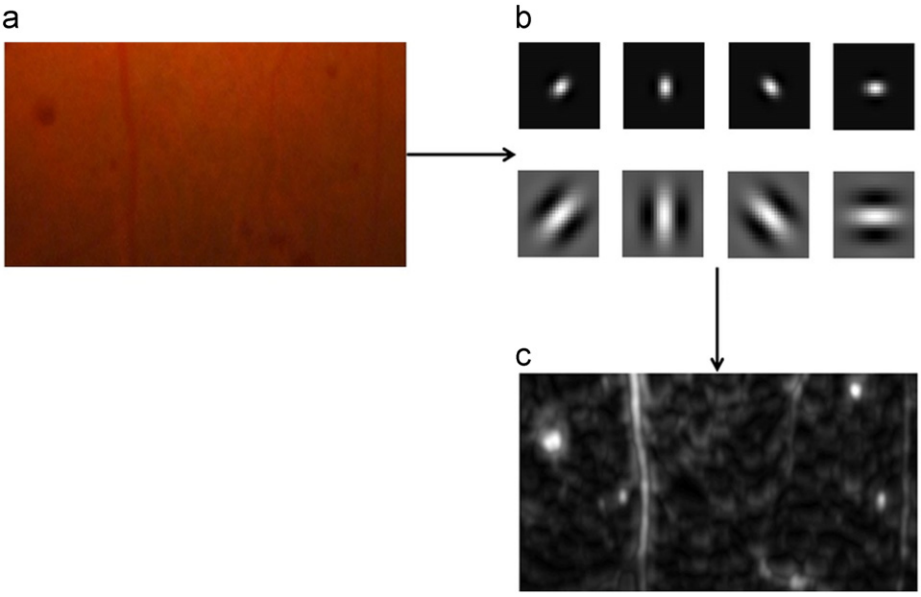

Fig. 4. Filter bank response: (a) slice from input image containing MAs; (b) total 8 Gabor filter bank kernels are used. First and second row contain kernels for scale = 2 and 5, respectively and with orientations = $\{\pi/4, \pi/2, 3\pi/4, \pi\}$; (c) maximum Gabor filter response after application of all eight kernels highlighting all MA regions.

For considered frequency and scale values, the maximum Gabor filter bank response $M_\gamma(\sigma,\Omega)$ is computed using Eq. (6) for θ spanning from 45° up to 180° at steps of 45°.

$$M_\gamma(\sigma,\Omega) = \max|\gamma(\sigma,\Omega,\theta)| \quad (6)$$

Fig. 4 shows the enhancement of regions with possible MAs using filter banks.

The binary candidate regions for MAs and *H* are extracted from M_γ by applying a low threshold value T_1. These regions also contain blood vessel pixels which are false lesion regions and need to be removed before further processing. In this system blood vessels are enhanced and segmented by using our previously proposed system given in [22]. The accurate segmentation of blood vessels and the removal of vessel pixels help in reducing the number of false positives. Fig. 5 shows the segmented candidate lesion pixels before and after the removal of blood vessel pixels.

5. Feature vector formation

MA regions appear with distinguishable properties such as color, size and shape. MAs are small in size and they appear in dark red colored circles shape. For an automated system to distinguish between MA and non-MA regions, a feature vector is formed for each candidate region. We have divided the features into four subsets such as

- Shape based features like area and shape of the candidate region.
- Gray level features based on gray level intensities inside the candidate region. These features are calculated using green channel of colored retinal image as it gives the best contrast between the fundus region and lesions present on it.
- Color features such as hue, saturation and value.
- Statistical features such as entropy, energy and moments.

Candidate region extraction stage gives all possible objects (regions) that can be considered as potential MA regions. If an image v contains N potential candidate regions, then the set representation for an image v is $v = \{v_1, v_2, v_3, \ldots, v_N\}$. Each object or candidate lesion region is considered as sample for classification and represented by a feature vector containing all m features, i.e. for a sample lesion v_i the feature vector is $v_i = \{f_1, f_2, f_3, \ldots f_m\}$, where $i = 1,2,\ldots,N$. The descriptions of features which we use in the proposed systems are as follows:

(1) *Area* is the total number of pixels in candidate region v_i.
(2) *Eccentricity* is the ratio of the distance between foci of ellipse and its major axis length and it is equal to 0 for a circular region.
(3) *Perimeter* is the boundary that surrounds the area of candidate region.
(4) *Compactness* $C = P^2/4\pi A$ is another measure of circularity, where A and P are the area and perimeter of the candidate region.
(5) *Aspect ratio* is the ratio of major axis length to minor axis length of the candidate region.
(6) *Mean and standard deviation* value of all green channel pixels within the candidate region.

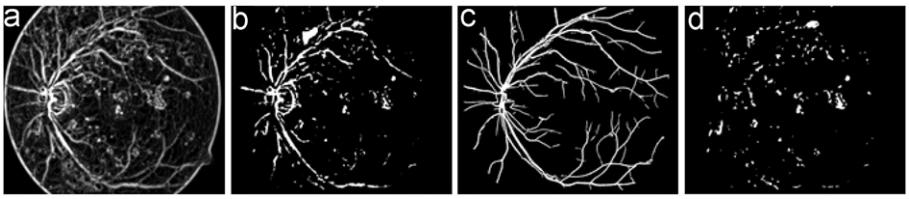

Fig. 5. Red candidate lesion enhancement and extraction. (a) Gabor filter banks response M_γ; (b) segmented red lesions containing spurious regions; (c) segmented blood vascular pattern using [22]; (d) candidate lesions after vessel subtraction.

(7) *Mean and standard deviation* value of contrast enhanced green channel pixels within the candidate region. The adaptive histogram equalization and Eq. (2) are used for contrast enhancement.
(8) *Mean gradient magnitude* value for boundary pixels of candidate lesion.
(9) *Mean gradient* value of neighbor pixels in a square region outside the candidate region.
(10) *Mean HSV (hue, saturation and value)* values of all pixels inside the candidate region.
(11) *Standard deviation HSV* values of all pixels inside the candidate region.
(12) *Entropy* value of all pixels in the square region including candidate region pixels and itsneighboring pixels.
(13) *Energy* value of all pixels in the square region including candidate region pixels and its neighboring pixels.
(14) *Homogeneity* value of all pixels in the square region including candidate region pixels and its neighboring pixels.
(15) *Third moment* value of all pixels in the square region including candidate region pixels and its neighboring pixels. It is the measure of skewness.

6. Classification using hybrid classifier

Given the feature space representation of candidate regions, we now describe our proposed hybrid classifier for the detection of MAs. We model the lesion classes (MAs and non-MAs) using a proposed ensemble of the state-of-the-art classifiers including GMM, SVM and an extension of multimodal *m*-Mediod based modeling approach, as presented in [29]. GMM and SVM classifiers are based on the work presented in [26,31], respectively. Multimodal *m*-Mediods approach is modified to act as an optimal classifier without considering the anomaly detection as it is not a requirement in lesion classification. Before generating *m*-Mediods based model of MA and non-MA classes, we perform feature preprocessing step to enhance separation between candidate region features belonging to MA and non-MA classes. We employ a supervised local Fisher discriminant analysis (LFDA) which achieves this by identifying principal directions with the maximized discrimination between different classes. This results in enhancing within class similarity and between class difference while catering for the presence of multimodal distribution of samples with a particular class. Let $DB=\{v_1,v_2,\ldots,v_n\}$ be a set of n training samples belonging to c classes. The within class and between class scatter matrix is computed as

$$\aleph_w=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}W_{i,j}^{w}(\|v_i,v_j\|) \tag{7}$$

$$\aleph_b=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}W_{i,j}^{b}(\|v_i,v_j\|) \tag{8}$$

where $\|.,.\|$ is a Euclidean distance function and

$$W_{i,j}^{w}=\begin{cases}\exp\left(\dfrac{\|v_i,v_j\|^2}{\varsigma_i\varsigma_j}\right)*\dfrac{1}{n_k} & \text{iff } v_i\wedge v_j\in\mathbf{C_k}\\ 0 & \text{otherwise}\end{cases} \tag{9}$$

$$W_{i,j}^{b}=\begin{cases}\exp\left(\dfrac{\|v_i,v_j\|^2}{\varsigma_i\varsigma_j}\right)*\left(\dfrac{1}{n}-\dfrac{1}{n_k}\right) & \text{iff } v_i\wedge v_j\in\mathbf{C_k}\\ \dfrac{1}{n} & \text{otherwise}\end{cases} \tag{10}$$

here n_k is the membership count of class $\mathbf{C_k}$ and ς_i is the average distance of sample v_i with its k nearest neighbors. This local scaling of affinity matrix caters for multimodal distribution of samples within a given class. The LFDA based coefficient space representation is generated by performing generalized eigenvalue decomposition of $\aleph_b E=\lambda\aleph_w E$, where λ is a generalized eigenvalue and E is the corresponding eigenvector. The enhanced feature space representation of candidate region, using localized Fisher discriminant directions, is then obtained as

$$F=\{E_1,E_2,\ldots,E_m\} \tag{11}$$

where $\{E_1,E_2,\ldots,E_m\}$ are eigenvectors arranged in descending order w.r.t. their corresponding eigenvalues $\{\lambda_1,\lambda_2,\ldots,\lambda_m\}$.

We now present the modified *m*-Mediods approach to adapt to the task of MAs detection.

6.1. Multimodel m-Mediods based modeling and classification

Given the LFDA-based enhanced feature space representation of regions, we generate the model of normality of MAs and non-MAs lesion classes. Our modeling approach, referred to as *m*-Mediods modeling, models the class containing n members with m mediods known *a priori* where each mediod is a cluster center of a mutually disjunctive sub-class within the lesion pattern. The algorithm for generating *m*-Mediods model is a Semi-Fuzzy Self-Organizing Map (SFSOM) based learning mechanism. Let $DB^{(i)}$ be the classified training samples associated to lesion class i and W be the weight vector associated to each output neuron. The modeling algorithm comprises the following steps:

(1) Initialize the SFSOM network with $\#_{output}$ neurons. We empirically set $\#_{output}=3*m$ upper bounded by the number of lesions present in a given lesion class (MAs or non-MAs).
(2) Initialize W_i (where $1\le i\le\#_{output}$) from the *PDF* $N(\mu,\Sigma)$ estimated from training samples in $DB^{(i)}$.
(3) Sequentially input feature vector representation of lesions F as training sample. Identify k Nearest Weights (k-NW) to current training sample:

$$k\text{-}NW\ (F,\mathbf{W},k)=\{P\in\mathbf{W}\,|\,\forall R\in P,S\in\mathbf{W}-C,\|F,R\|\le\|F,S\|\wedge|P|=k\} \tag{12}$$

where $\mathbf{W}$ is the set of all weight vectors and P is the set of k closest weight vectors. The value of k determines the

number of output neurons whose corresponding weight vectors are closest to F and will be updated in the corresponding learning iteration. For a given training cycle t, $k=\delta(t)$, where $\delta(t)$ is a neighborhood size function whose value decreases gradually over time as specified in Eq. (15).

(4) Train SFSOM network by updating a subset of the weights (P) using

$$W_c(t+1)=W_c(t)+\alpha(t)\zeta(j)(F-W_c(t)) \quad \forall W_c\in P \tag{13}$$

where W_c is the weight vector representation of output neuron c, j is the order of closeness of W_c to F ($1\le j\le k$), $\alpha(t)$ is the learning rate of SFSOM w.r.t. current learning cycle t and $\zeta(j,k)=\exp(-(j-1)^2/2k^2)$ is a membership function that has value 1 when $j=1$ and falls off with the increase in the value of j.

(5) Decrease the learning rate $\alpha(t)$ using

$$\alpha(t)=1-e^{(2(t-t_{max}))/t_{max}} \tag{14}$$

where t_{max} is the maximum number of training iterations lower bounded by the number of samples in training dataset.

(6) Decrease the neighborhood size exponentially over time

$$\delta(t)=\lceil\delta_{init}(1-e^{(2(t-t_{max}))/t_{max}})\rceil \tag{15}$$

where δ_{init} is the neighborhood size at the start of learning process. A series of experiments were conducted to determine the best value for δ_{init} which comes out to be 5.

(7) Repeat steps 3–6 for all the training iterations.

(8) Remove output neurons with zero membership count.

(9) Merge the selected pair of weight vectors using

$$W_{ab}=\frac{|W_a|\times W_a+|W_b|\times W_b}{|W_a|+|W_b|} \tag{16}$$

where $|.|$ is the membership count function and (a,b) is the closest pair of weight vectors given by the condition

$$(a,b)=\arg\min_{(i,j)}\|W_i,W_j\|\times\sqrt{|W_i|+|W_j|} \quad \forall i,j\wedge i\neq j \tag{17}$$

Scaling the distance between two weight vectors by their respective membership counts results in having more mediods presence in dense distribution of training samples belonging to a particular class and vice versa. This is crucial to the modeling of multimodal distribution of samples with a class.

(10) Iterate steps 8 and 9 till the number of weight vectors gets equivalent to $\#_{mediods}$. Append weight vector W_k to the list of mediods $\mathbf{M}^{(i)}$ modeling the pattern i.

(11) Compute the distance of each mediod from its k nearest mediods and calculate the mean to approximate the local density of the distribution. Append the mean distance to $\mathbf{D}^{(i)}$ in correspondence to a given mediod in the mediods list $\mathbf{M}^{(i)}$.

Once we have modeled the MAs and non-MAs classes using multivariate m-Mediods based model, the classification of unseen lesions is done by computing the relative closeness of feature vector representation of test lesion w.r.t. the models of different lesion classes. The proposed lesion classifier is comprised of following steps:

(1) Select k nearest mediods, from m-Mediod model corresponding to each class, w.r.t. feature vector representation of unseen lesion sample Q as

$$k\text{-}NM\{i\}(Q,\mathbf{M},k)=\{P\in\mathbf{M}^{(i)}\,|\,\forall R\in P,S\in\mathbf{M}^{(i)}-P,\ \|Q,R\|\le\|(Q,S\|\wedge|P|=k\} \quad \forall i \tag{18}$$

where $\mathbf{M}^{(i)}$ is the set of mediods from modeling class i and $\mathbf{P}$ is the ordered set of k closest mediods starting from the nearest mediod.

(2) Compute the fuzzy (probabilistic) membership $\Im\{i\}$ of test lesion sample from class i as

$$\Im\{i\}=1-\frac{\frac{\sum_{j=1}^{k}\|Q,k\text{-}NM\{i\}_j\|}{k}}{\overline{\mathbf{D}^{(i)}}} \tag{19}$$

where $\overline{\mathbf{D}^{(i)}}$ is the average of mean distances corresponding to mediods in $\mathbf{P}$ as identified in Eq. (18). The mean distance corresponding to a given mediod is precomputed and stored in $D^{(i)}$ as specified in step 11 of modeling algorithm. Classify sample to either MA or non-MA class having the highest probability ($\Im$).

6.2. Hybrid classifier

The proposed multimodal m-Mediod classifier is now combined with LFDA-GMM and SVM classifiers in a weighted probabilistic framework to obtain better decision. The classification of test lesion q using probabilistic classification prediction, based on the measure of evidence from different classifiers, is performed as

$$class(q)=\arg\max_{\forall class_i}\left(\sum_{i=1}^{c}a_k * P_{\mathbf{C_k}}(y=class_i|q)\right) \tag{20}$$

where $P_{\mathbf{C_k}}(y=class_i|q)$ is the probability of $class_i$ given a test lesion sample using classifier k and a_k is the weight associated to the probabilistic prediction of class $\mathbf{C_k}$. The weights are learned automatically as follows:

(1) Randomly divide labeled training data into different training and validation sets.

(2) Model training data using m-Mediods, SVM and LFDA-GMM based models.

(3) Classify test samples from corresponding validation sets using each of the generated model of normality.

(4) Filter those samples from validation set for which all the concerned classifiers give unanimous decision as these will not contribute in the weight learning process. Let $\mathbf{\Gamma}$ be a given validation set of lesion samples, the filtered validation set $\acute{\mathbf{\Gamma}}$ is obtained as:

$$\acute{\mathbf{\Gamma}}=\{\Gamma_i\in\mathbf{\Gamma}\,|\,class_{\mathbf{C}_j}(\Gamma_i)\neq class_{\mathbf{C_k}}(\Gamma_i)\ \forall j,k\wedge j\neq k\} \quad \forall i \tag{21}$$

where $class_{\mathbf{C}_j}(\Gamma_i)$ is the classification output of lesion sample Γ_i using classifier $\mathbf{C}_j$. This filtering of validation set significantly speeds up the weight learning process.

(5) Apply genetic algorithm to search for an optimal combination of weights which optimizes the accuracy of hybrid classifier using different validation sets randomly extracted from the classified training data.

7. Experiments

7.1. Material

The evaluation and testing of automated medical diagnostic systems are very important. We have used two publicly available retinal image databases namely DIARETDB0 [27] and DIARETDB1 [28] for evaluation purposes. The reason for choosing these datasets is that they contain a large variety of DR lesions and provide a good mean of evaluation. We formed three sets of images for testing using these two databases. First dataset (Set A) contains 130 retinal images, while the second (Set B) contains 89. Images were captured using the 50° Field of View (FOV) digital fundus camera with varying imaging settings and with resolution

of 1500×1152 pixels. The third dataset (Set C) is formed by combining Set A and Set B for more evaluation and testing purposes. These databases contain overall 219 retinal images, of which 25 images do not contain any sign of DR and are considered as normal retinal images. The remaining 194 images contain different signs of DR such as MAs, H, HE and CWS. The datasets also contain ground truths which are soft maps of DR lesions present in the image. A binary map is created for each image using the information given in ground truth for lesion level evaluation. We divided the complete dataset at lesion level into two subsets, i.e. training dataset and testing dataset. The detailed specifications of training and testing datasets are given in Table 1. It describes the dataset in terms of total number of lesions and training/testing.

7.2. Results

Here, we provide a quantitative evaluation and comparison of the proposed multivariate *m*-Mediods based classification algorithms with competitors such as GMM and one class classifier OCC-SVM. We also demonstrate the effectiveness of proposed approach to combine classifiers to enhance the classification accuracy of MA regions.

The experiment has been conducted on dataset1, dataset2 and dataset3. Feature vector representation of lesion from training data is generated as specified in Section 5 and are used to generate models as required by the different classification approaches. Multivariate *m*-Mediods based model of each class is generated using the algorithm as presented in Section 6. Patterns are modeled using 100 mediods per pattern. For SVM-based one class classifier (OCC-SVM), Tax et al. [31] performed classification by generating model of one class (referred to as target class) and distinguishing it from samples belonging to all other classes. There generation of model of the target class is done using SVM. We have employed radial basis function (RBF) kernel for the modeling of target class. Modeling of patterns and classification of unseen samples using GMM is based on the approach as described in [30]. Each class is modeled using a separate GMM. The number of modes to be used for GMM-based modeling is automatically estimated using a string of pruning, merging and mode-splitting processes as specified in [30]. Learning of weights for the proposed classifier ensemble, as specified in Eq. (20), is done using genetic algorithm based approach highlighted in procedure. Once the model learning phase is over, classification using different techniques is carried out using test dataset. We would expect that few instances drawn from class X would be classified to classes other than X and vice versa. The experiment is repeated with different combinations of training and testing samples. Each classification experiment is averaged over 50 runs to reduce any bias resulting from selection of a given combination of training and test samples. Fig. 6 shows examples of MAs detection. It shows the marked outputs of classification stage by eliminating the blood vessels and spurious region pixels.

The performance of proposed system is measured using sensitivity (Sen), specificity (Spec), positive predictive value (PPV) and accuracy (Acc) as figure of merit. These parameters are calculated using the following equations respectively:

$$\mathrm{Sen} = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FN})} \tag{22}$$

$$\mathrm{Spec} = \frac{\mathrm{TN}}{(\mathrm{TN}+\mathrm{FP})} \tag{23}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{(\mathrm{TP}+\mathrm{FP})} \tag{24}$$

$$\mathrm{Acc} = \frac{(\mathrm{TP}+\mathrm{TN})}{(\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN})} \tag{25}$$

where

- TP (True Positive): MA regions that are correctly classified by the classifier.
- FP (False Positive): Non-MA regions that are wrongly classified as MA regions by the classifier.
- TN (True Negative): Non-MA regions that are correctly classified by the classifier.
- FN (False Negative): MA regions that are wrongly classified as non-MA regions by the classifier.

To evaluate the performance of the proposed hybrid classifier (HC), the comparison of HC with GMM, SVM and m-Mediods (Med) is performed on all the three datasets separately. Tables 2–4 show the comparison of classifiers in the form of sensitivity, specificity and accuracy for set-A, set-B and set-C, respectively. We perform the evaluation at lesion level by comparing the outputs of classifier with the ground truths.

In order to perform further evaluation at feature level, we have used each descriptor, i.e. shape, gray, etc. separately to calculate the accuracy of each classifier. Fig. 7 shows the comparison of hybrid classifier with other classifiers using different types of descriptors. It is clear from figure that the effectiveness of hybrid classifier becomes prominent when we use each descriptor separately. The shape based features contribute the most in achieving high accuracy but still remaining descriptors in combination with shape based features help in getting even more higher accuracies. We have performed the Wilcoxon rank test to give a more insight about the effectiveness of each and individual feature. The Wilcoxon rank test is normally used for feature selection which tests whether the median of two classes differs significantly or not. Here, we are using this test just to highlight the contribution of each feature in overall accuracy. Table 5 shows the performance evaluation of features using the Wilcoxon rank test for MAs and non-MAs. The feature labels are according to the ones which are mentioned in feature formulation section. The features are arranged in descending order of their scores which mean features with high scores are at top and they contribute the most in accurate classification. Although the last few features can be excluded due to low scores, they do not have any diverse effect on accuracy so we did not perform feature selection here.

The classification accuracies obtained using different classifiers on different datasets are presented in Fig. 8. The results presented demonstrate the superiority of the proposed hybrid classifier as compared to LFDA-GMM, SVM and *m*-mediod. Another important observation from Fig. 8 is that combining classifiers using the proposed approach yields better accuracies as compared to individual classifiers.

We also perform a comparative analysis of our system with previously proposed techniques but the issue is that everyone has used their own local datasets and not all the databases are publicly available for bench marking. We have used the databases

Table 1
Datasets specifications.

Dataset	Total images	Total segments	Training segments	Testing segments
Set A	130	1876	938	938
Set B	89	1304	652	652
Set C	219	3180	1594	1586

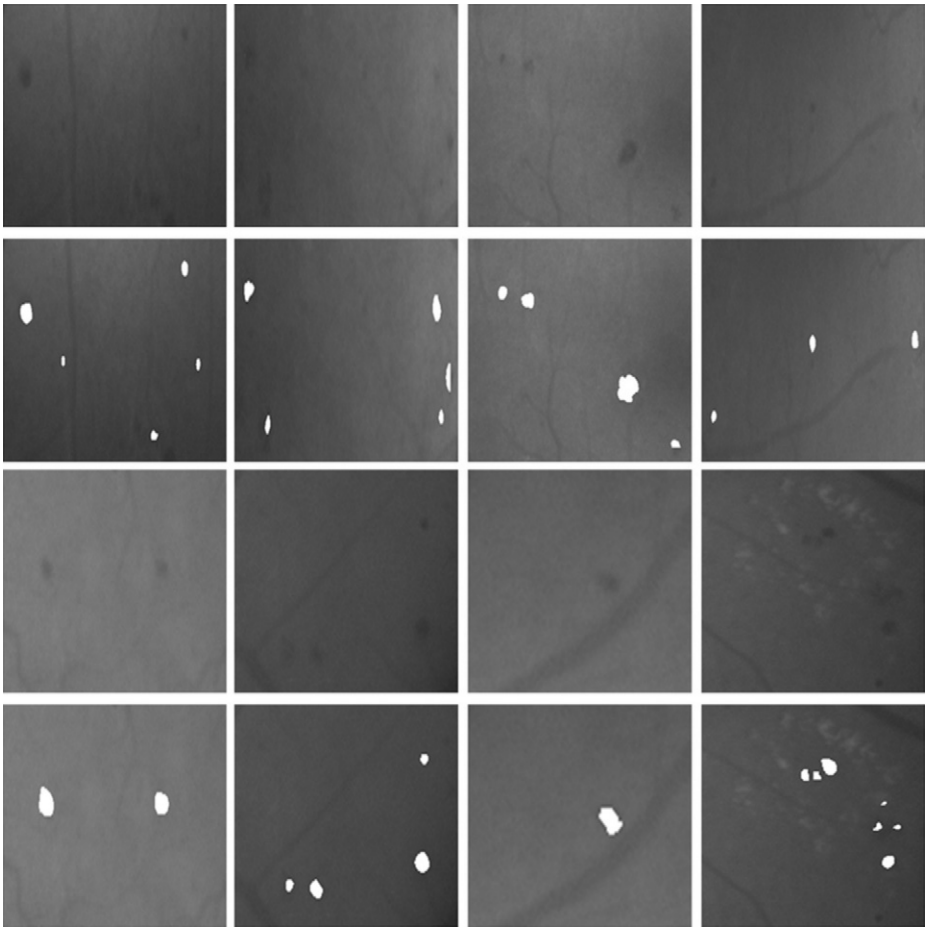

Fig. 6. MAs detection: different retina image slices containing MAs of variable structures. White dots are the regions classified as MAs using hybrid classifier.

Table 2
Performance evaluation comparison for MAs detection between GMM, SVM, m-Mediods (m-Med) and hybrid classifier (HC) for set-A.

Factors	GMM	SVM	m-Med	HC
TP	293	285	284	294
TN	638	637	640	641
FP	4	5	2	1
FN	3	11	12	2
Sen	98.98	96.28	95.94	99.32
Spec	99.37	99.22	99.68	99.84
PPV	98.65	98.27	99.30	99.66
Acc	99.25	98.29	98.50	99.68

Table 3
Performance evaluation comparison for MAs detection between GMM, SVM, m-Mediods(m-Med) and hybrid classifier(HC) for set-B.

Factors	GMM	SVM	m-Med	HC
TP	211	218	220	219
TN	432	430	427	430
FP	0	2	5	2
FN	9	2	0	1
Sen	95.90	99.09	100	99.54
Spec	100	99.53	98.84	99.53
PPV	100	97.75	97.77	99.09
Acc	98.61	99.38	99.23	99.53

Table 4
Performance evaluation comparison for MAs detection between GMM, SVM, m-Mediods(m-Med) and hybrid classifier(HC) for set-C.

Factors	GMM	SVM	m-Med	HC
TP	490	502	500	508
TN	1069	1068	1069	1070
FP	2	3	2	1
FN	25	13	15	7
Sen	95.14	97.47	97.08	98.64
Spec	99.81	99.71	99.81	99.90
PPV	99.59	99.40	99.60	99.80
Acc	98.29	98.99	98.92	99.49

that are publicly available for bench marking and comparison. Table 6 shows the comparison of different methods with proposed method (PM) for MA detection. It clearly highlights that the proposed system has outperformed other methods especially in case of specificity. This significance improvement is supported by accurate extraction of blood vessels which helped in removal of spurious regions and reduction of false positive regions.

In terms of complexity and execution time, the proposed hybrid classifier takes few milliseconds for processing each candidate lesion on a Core i-5 (2.3-GHz PC) with 4 GB of RAM. Table 7 shows the complexity and execution time of each classifier. Here, n are the number of classes and g, s and m are the number of Gaussian mixtures, support vectors and mediods per class, respectively. Normally, we require a more number of support vectors than mediods and Gaussian components resulting in higher computational time for SVM as compared to GMM and mediods. The complexity and execution time for hybrid classifier

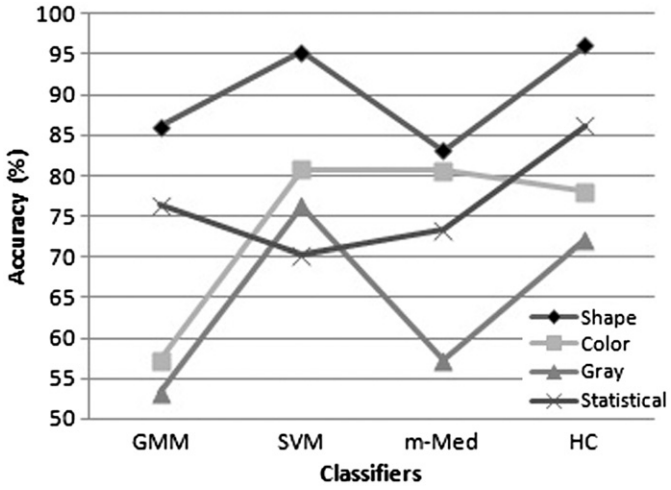

Fig. 7. Comparison of classifiers with different types of descriptors.

Table 5
Performance of all features calculated using the Wilcoxon rank test for MAs and non-MAs.

Feature	Descriptor	*Score*	*p*-Value
1	Shape	15	$<10^{-6}$
3	Shape	14.58	$<10^{-6}$
2	Shape	11.55	$<10^{-6}$
5	Shape	11.55	$<10^{-6}$
4	Shape	11.38	$<10^{-6}$
13	Gray	7.99	$<10^{-6}$
7	Statistic	5.21	$<10^{-6}$
10	Color	5.14	$<10^{-6}$
11	Color	4.94	$<10^{-6}$
15	Statistics	4.91	$<10^{-6}$
12	Statistics	3.71	0.00020
9	Statistics	2.81	0.00490
6	Gray	2.04	0.04146
14	Gray	2	0.04575
8	Gray	1.02	0.30766

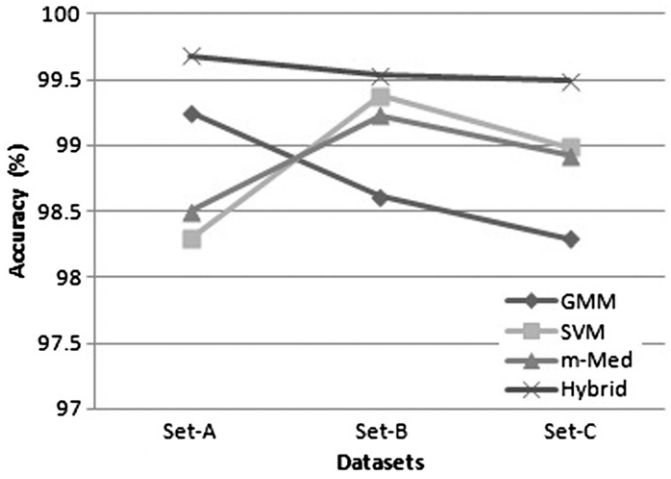

Fig. 8. Classification accuracies of classifiers using different datasets.

is the sum of all the three classifiers but still it is clearly evident that classification hardly takes anytime so its execution time is not very critical for system's perspective.

Table 6
Comparison of proposed method against previous techniques for MAs.

Methods	Sen	Spec	Acc
Niemeijer et al. [15]	100	87	–
Fleming [13]	85.4	83.1	–
Quellec [12]	89.62	–	0.929
Keerthi [8]	88.46	–	–
Walter [11]	88.5	–	–
Kahai [32]	100	67	–
Quellec [33]	–	–	0.927
Larsen [34]	71.4	96.7	82.6
Sinthanayothin [16]	77.5	88.7	–
PM	**98.64**	**99.69**	**99.40**

Table 7
Execution time and complexity analysis of different classifiers.

Classifier	Complexity	Execution time (s)
GMM	O(n*g)	2.93e−004
SVM	O(n*s)	8.28e−004
m-Mediods	O(n*m)	3.17e−004

8. Discussion and conclusion

This paper presented an automated system for accurate detection of MAs using colored retinal images. A classic three-stage model is used consisting of candidate region extraction, feature vector formation and classification. The candidate region extraction phase applied mathematical morphology, contrast enhancement technique and Gabor filter bank for the enhancement of MA regions present in fundus image. Blood vessel pixels are eliminated in candidate extraction phase to reduce the number of spurious regions. A detailed feature vector for each candidate region is formed consisting of four types of features such as gray level, colored, shape and statistical based features.

The feature space formed for each region is enhanced by applying a supervised LFDA. We applied three classifiers GMM, OCC-SVM and enhanced m-Mediods to check the accuracy of MA detection. In order to further improve the accuracy of classification, we proposed a hybrid classifier which is a weighted combination of all three classifiers using a genetic algorithm. Three datasets are used for proper evaluation of proposed technique at lesion level. We have used sensitivity, specificity, PPV and accuracy as merit of figure to evaluate the performance of our system. It is critical to know that the datasets which we have used contain all types of variations which can be in retinal images with respect to DR so that there is no ideal conditions are assumed.

The proposed three-stage system is in the form of chain process where accuracy of each stage depends on the accuracy of output from the previous stage. That is why the proposed system focused on improving the accuracy of each stage in order to have a higher accuracy overall. The accuracy of candidate region extraction stage is improved by extracting and eliminating all blood vessels pixels to reduce the existence of false positive regions. The feature vector consisted of accurate descriptors and the feature space is further enhanced using the supervised LFDA. Finally in order to improve the accuracy of classification stage, we proposed an extension in m-Mediods based classifier and used this in a hybrid classifier as a combination with GMM and SVM to ensure high accuracies. Although the accuracies achieved by hybrid classifier is not much different from other ones but still the improvement is there and this is what we need in

medical systems. Table 6 shows that the overall complexity and execution time of hybrid classifier is almost the same as compared to other classifiers so the improvement in accuracies made by the hybrid classifier has almost no extra effect on execution time. The results showed that the proposed system has classified MAs with high values of sensitivity, specificity, PPV and accuracy which makes the proposed system suitable for early detection of MAs.

References

[1] R. Klein, B.E.K. Klein, S.E. Moss, Visual impairment in diabetes, Ophthalmology 91 (1994) 1–9.

[2] A.K. Sjolie, J. Stephenson, S. Aldington, E. Kohner, H. Janka, L. Stevens, J. Fuller, EURODIAB Complications Study Group, Retinopathy and vision loss in insulin-dependent diabetes in Europe, Ophthalmology 104 (1997) 252–260.

[3] Effective Health Care—Complications of Diabetes: Screening for Retinopathy and Management of Foot Ulcers, Royal Society of Medicine Press 5 (4) (1999), ISSN 0965-0288.

[4] P.C. Ronald, T.K. Peng, A Textbook of Clinical Ophthalmology: A Practical Guide to Disorders of the Eyes and Their Management, 3rd ed., World Scientific Publishing Company, Singapore, 2003.

[5] R.N. Frank, Diabetic retinopathy, Progress in Retinal and Eye Research 14 (2) (1995) 361–392.

[6] E.J. Susman, W.J. Tsiaras, K.A. Soper, Diagnosis of diabetic eye disease, Journal of the American Medical Association 247 (23) (1982) 3231–3234.

[7] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Automated identification of diabetic retinal exudates in digital colour images, British Journal of Ophthalmology 87 (10) (2003) 1220–1223.

[8] K. Ram, G.D. Joshi, J. Sivaswamy, A successive clutter-rejection-based approach for early detection of diabetic retinopathy, IEEE Transactions on Biomedical Engineering 58 (3) (2011) 664–673.

[9] B. Antel, A. Hajdu, Improving microaneurysm detection using an optimally selected subset of candidate extractors and preprocessing methods, Pattern Recognition 45 (1) (2012) 264–270.

[10] B. Zhang, X. Wu, J. Yo, Q. Li, F. Karray, Detection of microaneurysms using multi-scale correlation coefficients, Pattern Recognition 43 (6) (2010) 2237–2248.

[11] T. Walter, P. Massin, A. Erginay, R. Ordonez, C. Jeulin, J. Klein, Automatic detection of microaneurysms in color fundus images, Medical Image Analysis 11 (6) (2007) 555–566.

[12] G. Quellec, M. Lamard, P.M. Josselin, G. Cazuguel, B. Cochener, C. Roux, Optimal wavelet transform for the detection of microaneurysms in retina photographs, IEEE Transactions on Medical Imaging 27 (9) (2008) 1230–1241.

[13] A.D. Fleming, S. Philip, K.A. Goatman, J.A. Olson, P.F. Sharp, Automated microaneurysm detection using local contrast normalization and local vessel detection, IEEE Transactions on Medical Imaging 25 (9) (2006) 1223–1232.

[14] Retinopathy Online Challenge Web Site. The University of Iowa and the ROC Organizers (Online). Available from: 〈http://roc.healthcare.uiowa.edu/results.php〉.

[15] M. Niemeijer, B.V. Ginneken, M. Cree, A. Mizutani, G. Quellec, C.I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A.M.Q. Li, Y. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, M. Abramoff, Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs, IEEE Transactions on Medical Imaging 1 (29) (2010) 185–195.

[16] C. Sinthanayothin, J.F. Boyce, T.H. Williamson, H.K. Cook, E. Mensah, S. Lal, D. Usher, Automated detection of diabetic retinopathy on digital fundus images, Diabetic Medicine 19 (2) (2002) 105–112.

[17] D. Usher, M. Dumsky, M. Himaga, T. Williamson, S. Nussey, J. Boyce, Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening, Diabetic Medicine 21 (2004) 84–90.

[18] Y. Hatanaka, T. Nakagawa, Y. Hayashi, M. Kakogawa, A. Sawada, K. Kawase, T. Hara, H. Fujita, Improvement of Automatic Hemorrhages Detection Methods using Brightness Correction on Fundus Images, in: Proceedings of the SPIE Medical Imaging, vol. 6915, 2008.

[19] U.R. Acharya, K.C. Chua, E.Y.K. Ng, W. Wei, C. Chee, Application of higher order spectra for the identification of diabetes retinopathy stages, Journal of Medical Systems 32 (6) (2008) 431–488.

[20] S.C. Lee, E.T. Lee, Y. Wang, R. Klein, R.M. Kingsley, A. Warn, Computer classification of nonproliferative diabetic retinopathy, Archives of Ophthalmology 123 (6) (2005) 759–764.

[21] Harihar Narasimha-Iyer, Ali Can, Bandrinath Roysam, Charles V. Stewart, Howard L. Tanenbau, Anna Majerovics, Hanumant Singh, Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy, IEEE Transactions on Biomedical Engineering 53 (6) (2006) 1084–1098.

[22] M.U. Akram, S.A. Khan, Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy, Engineering with Computers, 2011, doi: 10.1007/s00366-011-0253-7.

[23] T. Walter, J.C. Klein, P. Massin, A. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy—detection of exudates in color fundus images of the human, IEEE Transactions on Medical Imaging 21 (10) (2002) 1236–1243.

[24] C. Sinthanayothin, J.F. Boyce, H.L. Cook, T.H. Williamson, Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images, British Journal of Ophthalmology 83 (1999) 902–910.

[25] J. Sung, S.Y. Bang, S. Choi, A Bayesian network classifier and hierarchical gabor features for handwritten numeral recognition, Pattern Recognition Letters 27 (1) (2006) 66–75.

[26] J. Lu, Localized Fisher discriminant analysis based complex chemical process monitoring, AIChE Journal 57 (7) (2011) 1817–1828. 〈http://onlinelibrary.wiley.com/doi/10.1002/aic.v57.7/issuetoc〉.

[27] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, J. Pietilä, DIARETDB0: Evaluation Database and Methodology for Diabetic Retinopathy Algorithms, Technical Report, 2005.

[28] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, J. Pietilä, DIARETDB1 Diabetic Retinopathy Database and Evaluation Protocol, Technical Report, 2006.

[29] S. Khalid, Activity classification and anomaly detection using m-Mediods based modeling of motion patterns, Pattern Recognition 43 (10) (2010) 3636–3647.

[30] F.I. Bashir, A.A. Khokhar, D. Schonfield, Automatic object trajectory based motion recognition using Gaussian mixture models, in: IEEE International Conference on Multimedia and Expo, Netherlands, 2005, pp. 1532–1535.

[31] D.M.J. Tax, One-Class Classification, Ph.D. Thesis, Delft University of Technology, 2001, 〈http://ict.ewi.tudelft.nl/davidt/thesis.pdf〉.

[32] P. Kahai, K.R. Namuduri, H. Thompson, A decision support framework for automated screening of diabetic retinopathy, International Journal of Biomedical Imaging (2006) 1–8.

[33] G. Quellec, Stephen R. Russell, Michael D. Abràmoff, Optimal filter framework for automated, instantaneous detection of lesions in retinal images, IEEE Transactions on Medical Imaging 30 (2) (2011) 523–533.

[34] M. Larsen, J. Godt, N. Larsen, H. Lund-Andersen, A.K. Sjolie, E. Agardh, H. Kalm, M. Grunkin, D.R. Owens, Automated detection of fundus photographic red lesions in diabetic retinopathy, Investigative Ophthalmology and Visual Science 44 (2) (2003) 761–766.

Usman Akram is currently a Ph.D. student in Department of Computer Engineering at College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan. His research interests include biomedical engineering, image processing, machine learning and pattern recognition.

Shehzad Khalid graduated from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan in 2000. He received the M.Sc. degree from National University of Science and Technology, Pakistan in 2003 and the Ph.D. degree from the University of Manchester, UK in 2009. He is currently an Assistant Professor at the Bahria University of Management and Computer Sciences, Pakistan. His research interests includes dimensionality reduction, indexing and retrieval, profiling and classification, trajectory-based motion learning profiling and classification, computer vision and machine learning.

Shoab A. Khan had a Ph.D. degree in Computer Engineering from Georgia Tech, USA. He is currently the head of computer engineering department at College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan. He has worked on various projects in the field of embedded systems, communication, biometrics, signal processing etc.