



Detection of Hard Exudates Using Evolutionary Feature Selection in Retinal Fundus Images

Anoop Balakrishnan Kadan¹ · Perumal Sankar Subbian²

Received: 12 March 2019 / Accepted: 20 May 2019 / Published online: 29 May 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

It is one of the most vital symptoms of DR (diabetic retinopathy) called hard exudates (HE), which are the leakage of cellular debris and lipoprotein from damaged blood vessels of retina. The vision loss is avoided if the detection of HE in the beginning times. Therefore, a novel method is proposed to detect hard exudates automatically. Previously, for exudate prediction supervised and unsupervised methods have been used. Fault detection of hard exudates, miss classification rate will affect these models because of the characteristics like, similarities with other components in the retinal image and intra variations. For that, the retinal fundus images has been used as input. Then these images are pre-processed with some pre-processing algorithms like image enhancement, equalization of histogram to improve the proposed system performance. Total image data files are divided to training and testing datasets. Features are extracted for training and testing using feature extraction algorithm individually. Then classifier algorithm predicts whether the hard exudate is proliferative or non-proliferative. We obtained accuracy of 99.34% using our proposed methods on public datasets like DIARETDB1 and DRIVE.

Keywords Fundus images · Diabetic retinopathy · Hard exudates · Proliferative · Histogram · Feature selection

Introduction

The most internal membrane of the human eye is retina. The eye sight affected directly if the diseases occur in the retina [1]. Hypertension and diabetes mellitus are the systemic disease in eye which cause some pathological changes. Pathological information can be provided by fundus eye digital images [2].

The grown countries consider diabetic retinopathy as a major causes of blindness and vision defects [3]. Fundus images permits retinal fundus high quality record for diagnosing DR early

signs and monitoring its progression. Automatic analysis is allowed by digital nature which minimize the ophthalmologist's workload as well as the cost of health in the disease screening [4]. This DR disease will grow from NPDR to PDR [5].

The NPDR is the detection of DR in early stage. The NPDR does not affect the vision and which begins with a mild NPDR [6]. The exudates are fluid leakage in blood which occur if the vision is affected. Hard exudates and Soft exudates are the different classification of exudates. The hard exudates are yellow spots in the retina and the soft exudates are look like white area with ill-defined edges or pale yellow [7]. Automated or computer aided analysis are required for screening programs for DR of retinal images [8]. The main normal features are blood vessels, Optic disk, macula and fovea [9]. The abnormal features in color fundus images include neovascularisation, exudates, hemorrhages, microaneurysms and cotton wool spots [10].

The abnormal retinal image classification is done with feature extraction. But most of the extracted features are redundant and completely irrelevant for classification target [11]. The sufficient data from features are not known in the initial stage to discriminate over the classes. The optic nerve head (ONH) is the important feature of DR [12]. The OD appears in the left or right side of the fundus images. The other landmarks

This article is part of the Topical Collection on *Image & Signal Processing*

✉ Anoop Balakrishnan Kadan
anoopbk2000@gmail.com

Perumal Sankar Subbian
spsankar2004@gmail.com

¹ Department of Electronics and Communication Engineering, Vimal Jyothi Engineering College, Chemperi Kannur, Kerala 670632, India

² Department of Electronics and Communication Engineering, Toc H Institute of Science and Technology, Arakkunnam, Ernakulam, Kerala 682313, India

of fundus images such as macula and fovea are located with the help of OD [13]. OD appears as a white region or bright yellowish in color fundus images. Outgoing vessels interrupts its shape [14].

There is no possibility to comprise all features needed for patterns or object classification. Feature selection is considered as the major task in classification problems [15]. It must find the number of features used in the classification as well as to maintain Suitable classification accuracy [16]. There are several machine learning procedures are presented for detection and prediction of exudates. [17].

There are number of authors discussed about normal/abnormal detection of retinal images recently. We have analyse the previous techniques and introduce an efficient method to detect the proliferative/non-proliferative hard exudates from abnormal retinal images. Most of the previous works developed on the basis of multi-layer perceptron (MLPs), neural networks (NN), Multi-scale texture classification, bank-note classification and support vector machines (SVMs). To overcome these diverse issues and applications, various segmentation techniques have been previously suggested to discriminate its accuracy but have not up to the mark.

By considering all the challenges listed above, we have proposed an efficient technique named as Enhanced Gaussian Mixture Model (EGMM) and KNN (K-Nearest Neighbour) classifier algorithm. Here, the pre-processing step is carried out initially consisting of gray conversion, denoising and image enhancement. Pre-processing is used to enhance and noise-reducing process for input retinal image. The main features such as blood vessels, macula, fovea, and exudates are extracted. Once the features are computed, training of KNN is done to classify the images into its proliferative or non-proliferative hard exudates from abnormal retina images.

This research paper is structured into six sections. Section 1 contains the introduction about diabetic retinopathy. In Section 2 we discussed the papers which are related to our work. In Section 3, we presented our proposed work which includes feature extraction, selection, dimension reduction, and exudate classification. Section 4 contains the experimental analysis which includes dataset description, performance metrics and performance analysis. Section 5 describes the conclusion and Section 6 contains the references.

Related work

Some of the recent research works related to the exudate prediction are listed below

Carla Pereira et al. [18] proposed an unsupervised approach which was based on an ant colony optimization for exudates segmentation. Performance of the algorithms was evaluated with an available online dataset and the result of the

experiment showed better results than the existing kirsch filter in exudate detection. Intelligent computer algorithms are needed for making decision for large volume of data screening.

T. Jaya et al. [19] proposed a FSVM (Fuzzy Support Vector Machine) classifier as a decision making system in fundus images for hard exudates detection. Here, the optic discs are segmented using circular Hough transform and morphological operations to avoid false alarms. Texture and color features were extracted to separate the exudates from the non-exudates pixels from the input. These extracted features were fed as an input to the FSVM classifier. 200 retinal images were analysed by the classifier and the inputs were collected from diabetic retinopathy screening programmes.

Sundararaj Wilfred Franklin et al. [20] proposed an algorithm based on neural network to automatically detect the exudates presence which helps the ophthalmologists for the diagnosis and DR follow-up. Colour, shape, texture and size features were considered in this work. Harry Pratt et al. [21] identified the complicated features using a CNN architecture and data augmentation in the classification task such as haemorrhages on the retina, micro-aneurysms and automatic diagnosis of user input. The publicly available kaggle dataset was used for train the network with the help of graphics processor unit (GPU) and which demonstrated impressive outcomes.

Usman Akram et al. [22] proposed a novel hybrid classifier system for retinal lesions detection. The hybrid classification consists of pre-processing, candidate lesion extraction, formulation of feature set and classification. The background pixels were eliminated by the system and extraction of blood vessels and optic disc were done in pre-processing. Different type of features were formulated to classify the region.

An extension of Gaussian Mixture Model is combined for hybrid classifier which always improves the classification accuracy.

Methodology for exudates prediction

The aim of this paper is hard exudates identification from retinal fundus images with high accuracy. The detection accuracies relied on the selection of feature extraction and machine learning algorithms. Five steps are used and discussed in this sections such as pre-processing, feature extraction, feature selection and classification using KNN. The proposed methodology in block diagram is illustrated in Fig. 1. It shows the classification of two types hard exudates such as proliferative and non-proliferative. Training and testing images are extracted from the retinal fundus images and pre-processing done for both images. Then feature extraction and optimal feature selection are done and these selected features are fed to KNN classifier for hard exudate prediction.

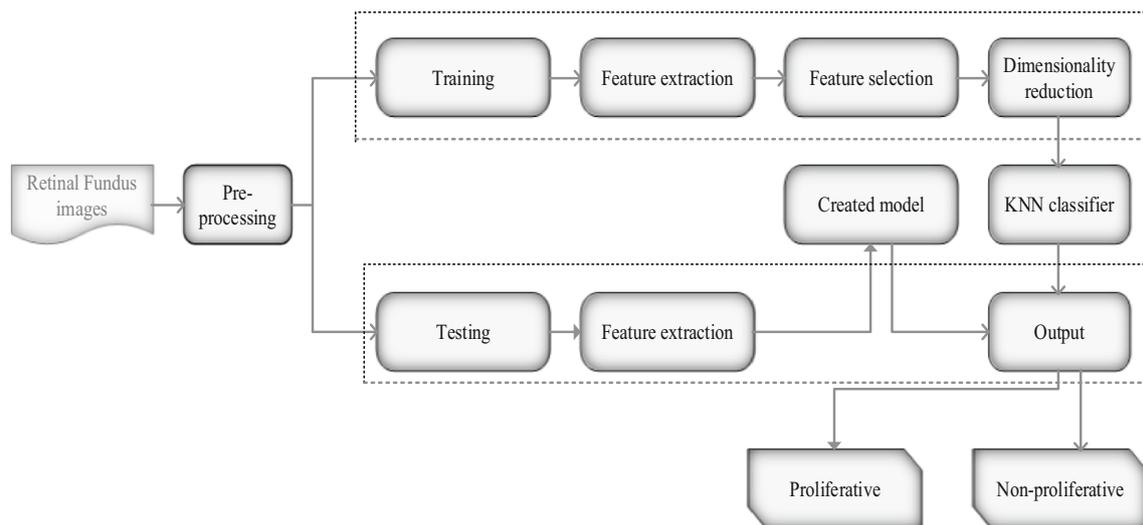


Fig. 1 Schematic representation of proposed methodology

For robust feature extraction, histogram approach is applied. Then some statistical features are computed to find out the hard exudate portion. Then the optic disk is removed by applying morphological operations like erosion and dilution.

The classification challenges are rectified with the help of feature selection. Gray wolf optimization is applied in feature selection of proposed approach and the dimension is reduced with fisher criterion. Finally KNN (K-Nearest Neighbour) was used as the learning method for exudate classifier. The implementation of proposed work is done with MATLAB tool and performance comparison is done with existing classifiers such as CNN,SVM and ANN.

Pre-processing stage

It is required for ensuring that the dataset is consistent and displays only relevant features. This step is necessary to simplify the workload of the following processes.

Channel separation

RGB Channel separation is the method that read the RGB channels separately. The image is denoted as $I(q_1, q_2, q_3)$ where $q_1 \in \{0, \dots, P - 1\}$, $q_2 \in \{0, \dots, Q - 1\}$ and $q_3 \in \{0, 1, 2\}$ is isolated into three gray scale images, $\{I_R(q_1, q_2), I_G(q_1, q_2), I_B(q_1, q_2)\}$. Where, (q_1, q_2) are the spatial dimension in a 2D image. $\{I_R(q_1, q_2), I_G(q_1, q_2), I_B(q_1, q_2)\}$ Represents the Red, Green, Blue color values of the image at q_1, q_2 pixel respectively. Splitting an image in its color channels also decreases the time complexity of algorithms. After converting the RGB to gray conversion, the gray image is moved to denoising.

Denoising

Adaptive median filter adjusts the filter window size according to the noise range of window centre. If the pixel of filter window center is the noise point, the median value is used to replace the pixel. If the pixel of filter window center is not noise point, its current value remains unchanged. Adaptive median filter can deal with impulse noise with larger noise intensity. Meanwhile, it can maintain the image details very well.

The detailed steps of adaptive median filter are as follows, where q_{min} and q_{max} are the minimum and maximum value of filter window. The median value in the filter window is denoted as q_{med} , $q_{i, j}$, is the gray value in $((i, j))$; and S_{max} is the maximum that is specified.

- Step 1: (Calculate) $Z_1 = x_{med} - x_{min}$ and $z_2 = x_{med} - x_{max}$
- Step 2: If $z_1 > 0$ and $z_2 < 0$ go to Step 3; otherwise, we extend the window $S_{i, j}$. If $S_{i, j} < S_{max}$, repeat Step 1 and Step 2. If not, output $x_{i, j}$, directly.
- Step 3: Calculate $k_1 = x_{ij} - x_{min}$ and $k_2 = x_{ij} - x_{max}$.
- Step 4: If $k_1 > 0$ and $k_2 < 0$, output $x_{i, j}$, directly. Otherwise we regard x_{med} as the output value.

The feature of this algorithm is that is that it can filter impulse noise and smooth non-impulse noise, besides it can reduce image distortion and protect image details at the same time. The denoised image is given as an input of background subtraction.

Background subtraction

Various computer vision approaches are designed with the help of background subtraction. Enhanced Gaussian Mixture Model (EGMM) is one of the best background subtraction

method in which the pixels are modelled into Gaussian distribution. Initially all the pixels are separated by its intensity in RGB color space. For probability computation, every pixels are used whether it is background or foreground with the following condition:

$$P(q_t) = \sum_{i=1}^k w_{i,t} \eta(q_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

Where, the current pixel in frame t is denoted as q_t and distributions in the mixture count is denoted as K . The k th distribution weight and mean in frame t is denoted as $w_{i,t}$ and $\mu_{i,t}$. $\Sigma_{i,t}$ represents k th distribution the standard deviation in frame t . The probability density function is denoted as $\eta(q_t, \mu_{i,t}, \Sigma_{i,t})$.

$$\eta(q_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp^{-1/2(q_t - \mu)^T \Sigma^{-1} (q_t - \mu)} \quad (2)$$

Each RGB is uncorrelated one with another one. So the intensity difference could be assumed possessing uniform standard deviation. The formulation of covariance matrix is denoted by:

$$\Sigma_{i,t} = \sigma_{i,t}^2 I \quad (3)$$

In Gaussian process, if the threshold is higher than the designated threshold considered as background and the remaining is called foreground.

$$B = \operatorname{argmin}_b \left(\sum_{i=1}^b w_{i,t} > T \right) \quad (4)$$

The value of w , μ and σ is updated when the pixel equals with one of the K Gaussian.

$$w_{i,t+1} = (1-\alpha)w_{i,t} + \alpha \quad (5)$$

$$\mu_{i,t+1} = (1-\rho)\mu_{i,t} + \rho \cdot q_{t+1} \quad (6)$$

$$\sigma_{i,t+1}^2 = (1-\rho)\sigma_{i,t}^2 + \rho(q_{t+1} - \mu_{i,t+1}) \cdot (q_{t+1} - \mu_{i,t+1})^T \quad (7)$$

Where,

$$\rho = \alpha \times \eta(q_{t+1}, \mu_i, \Sigma_i) \quad (8)$$

In the meantime, only the 'w' is updated, if there is a case where all K of Gaussian do not equal then

$$w_{j,t+1} = (1-\alpha)w_{j,t} \quad (9)$$

The foreground identification can be done if every parameter has been found.

Contrast enhancement

The intensity values of an image is transformed in this process such that the output image histogram equals the stated histogram approximately. The global contrast of the image is maximized by this contrast enhancement method. The local contrast of image is enhanced with adaptive histogram equalization. It is considered as a good tool for edge enhancement. After completing this process the enhanced image is moved to optic disc removal.

Removal of optic disc

The optic disc has few attributes like hard exudates: sharp boundaries and bright intensities. We initially identify the optic disc and remove it from attention to protect the disc from intrusive with exudate detection. Generally, optic disc has some attributes which are constant size, circular shape and high intensity and these attributes are exactly indicate the exudate location. In the optic disc there is some dark objects present which are wrongly noted as MAs (Microaneurysms) or HAs (Hard Exudates). Therefore, optic disc removal is important for avoiding the above mentioned problem. For this purpose only we starts with optic disc removal with concentrating on the mid third of the green intensity image FG as a ROI.

Feature extraction

The most important features of detecting hard exudates are texture, shape, color and size. One of the shape extraction technique is thresholding in which the images are viewed as the result to separate the user eye from the background. Thresholding method generates uniformity regions within an image on the basis of some threshold criterion T which is defined as given below,

$$T = T\{q_1, q_2, A(q_1, q_2), f(q_1, q_2)\} \quad (10)$$

Where, $f(q_1, q_2)$ is the gray level pixel at (q_1, q_2) and $A(q_1, q_2)$ defined as local property in the pixel neighbourhood. A threshold image is represented as,

$$g(q_1, q_2) = 1 \text{ if } f(q_1, q_2) \geq T \quad (11)$$

$$g(q_1, q_2) = 0 \text{ if } f(q_1, q_2) < T \quad (12)$$

The input image is partitioned into sub images in local thresholding approach which defines the threshold for all sub images. Normally, macula, fovea, exudates and blood vessels are presented in the retinal images. These features are extracted by applying thresholding concept. From the

exudates, 19 more statistical features are extracted to improve the classification accuracy such as contrast, energy, entropy, auto correlation, and sum average and sum variance. The other features are optic disk, blood vessel extraction, exudates and fovea.

Fovea

Fovea is presented in the darkest region of macula which is represented in Fig. 2. Fovea detection is very difficult than OD because there is no specific region like OD and its size is very small. Fovea relative position regarding OD location is very important. The estimated center of fovea is obtained by applying some morphological operations and computations.

Macula

In Fig. 2 the square box represent the macula and the black dot represents the fovea.

Blood vessels

For earlier diagnosis of disease, Blood vessel segmentation is accomplished. The important indicators for diseases are the physical changes the retinal blood thus, correct and exact segmentation of blood vessel is of critical significance. Blood vessels of the retinal fundus image is represented in Fig. 3.



Fig. 2 Macula and fovea

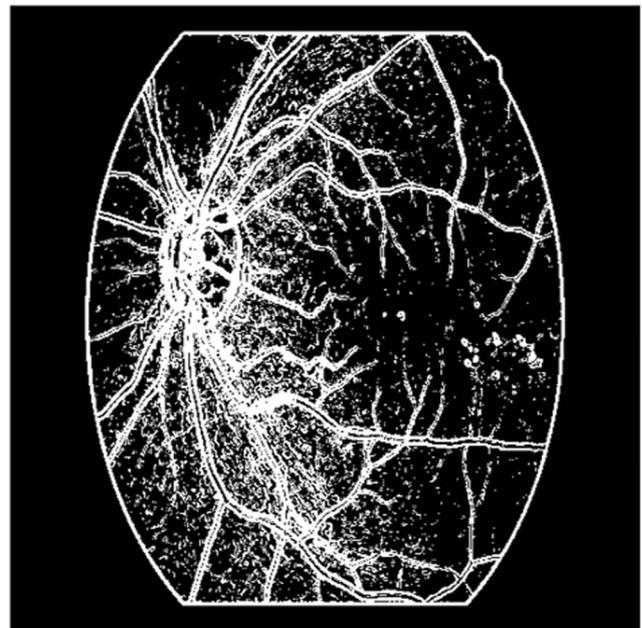


Fig. 3 Blood vessels

Exudates

With the help of morphological processing, the exudate regions are extracted. For the reason of high contrast between the vessels and background, in this scheme compulsory. In gray scale morphological closing operation, this process to eradicate the vascular system by having altered structured component. The bright objects are attained by using the threshold.

Feature selection

The number of attributes are reducing in this section. It selects a subset of original features. In data pre-processing, frequently used technique is a feature selection is used to classify the significant features. The selected features are often unknown previous and removes redundant or irrelevant features which do not have significance in classification task. To select the subset of valuable features, Grey wolf optimization algorithm is utilized in this work.

Grey wolf optimization

For leadership hierarchy and hunting mechanism of grey wolves, the Grey Wolf Optimizer (GWO) [23] is used. Based on the social hierarchy and attack, the hunting behaviour of grey wolf algorithm are searching the victim and encircling the victim. For optimal feature selection issue, the moralities of grey wolf optimization are used. In such a space, each feature subset can be seen as a location. If there are N total features, then there will be 2^N dissimilar feature subset

which is diverse from each other based on the length and features are encompassed in each subset. With least length, the optimal location is the subset and obtained more classification accuracy. For choosing the optimal feature set, the grey wolf optimization is applied. Finally, they should congregate on best, possibly optimal, positions. In the feature space, the GWO creates repetitions of investigation of different regions and exploiting solution until getting near-optimal solution. Figure 4 represents the flowchart for proposed hard exudate prediction model.

Dimensionality reduction

For the pattern recognition, in data dimension there are always various problems, when utilizing the statistical techniques. It is difficult to apply the methods in high-dimensional space which are worked in the low-dimensional space. Generally, problems in the low-dimensions can be solved with less

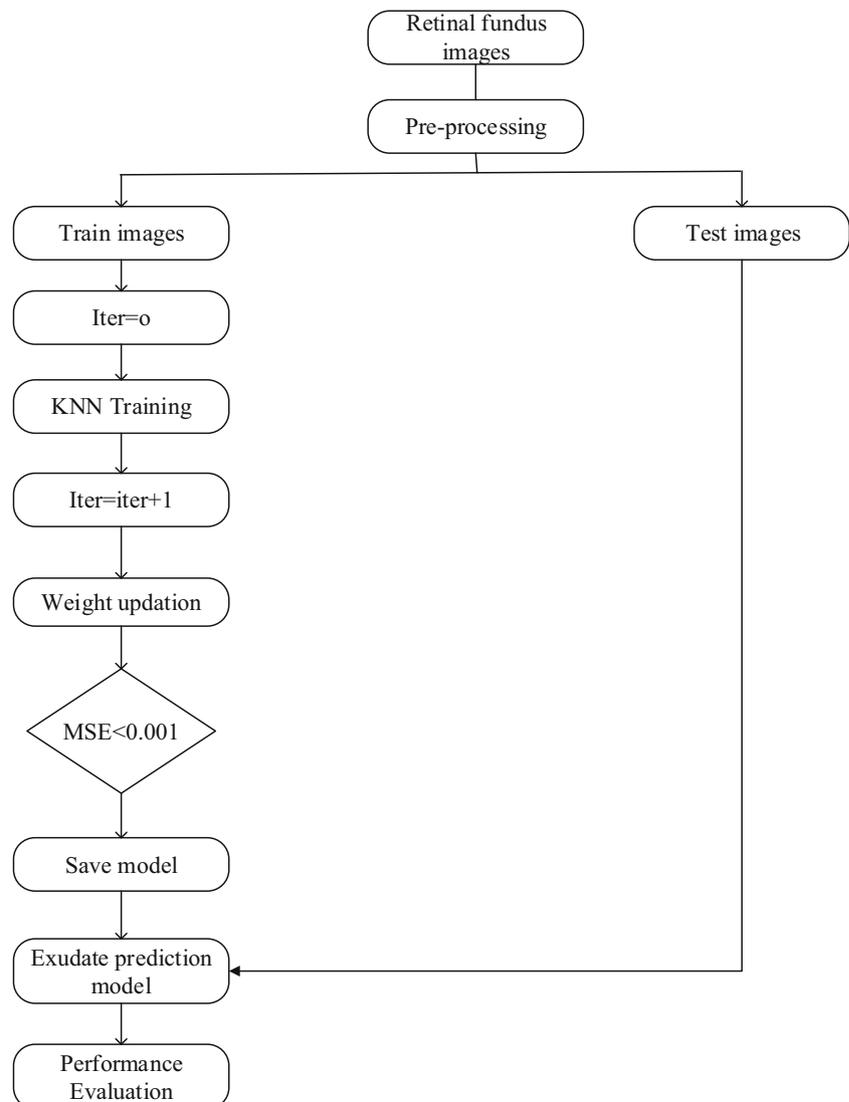
convenient, efficiency and computational complexity. Once the features are attained by the correlation analysis, a set of high-dimensional features which represent the extracted data sets are utilized for the testing and training of pattern classifiers. The main issue of the dimensionality is removing the irrelevant features. For that problem a pre-processing method is required for the classifier.

Fisher criterion and principal component analysis (PCA) are the well-known techniques for reducing the dimensions of the features. But the Fisher criterion method has the capability to solve the problems which are not capable to solve by the PCA for extracting the high-dimensional features information.

Fisher criterion

The fisher rate of every component computation is done for the ability to distinguish exudates. Then the value of ratios are

Fig. 4 Flowchart of proposed model



compared to find the optimal features. The representation of fisher criterion [24] is given below:

Step 1.

$$\psi_{Fisher} = \frac{\omega_{between}}{\omega_{within}} \tag{13}$$

The feature component fisher rate is defined as ψ_{Fisher} and the variance between the feature component classes is denoted as $\omega_{between}$. The representation of ω_{within} is variance inside features within the same component of the variance. The computation of $\omega_{between}$, ω_{within} are done by the following equations.

$$\omega_{between} = \sum_{i=1}^M \left(m_k^{(i)} - m_k \right)^2 \tag{14}$$

$$\omega_{within} = \sum_{i=1}^M \frac{1}{n_i} \sum_{j=1, c \in \omega_i} \left(c_k^{(ij)} - m_k^{(i)} \right)^2 \tag{15}$$

Where, i is the specific class symbol and M is the feature class count. The feature dimension is denoted as k . m_k signifies all samples mean of the k dimension. n_i denotes the number of samples for class i and sample symbol represented as j . The i class feature sample gather represented as ω_i .

The fisher criterion procedure is designed with the feature class specification.

- 1) Criterion arithmetic gets the input of feature class.
- 2) Each dimension fisher rate is computed based on Fisher criterion arithmetic.
- 3) The Fisher rate class summation is computed.
- 4) The S feature conferring to Fisher from big to small and choose the S sequence to T based on the sequenced order until $\psi_{Fisher} < \frac{1}{n_i} SUM$, where ‘ n ’ is the S element count.
- 5) Obtain the feature class T after selection.

Hard exudate classification

KNN classifier directly predicts the class using the dataset which is used for training. A new instance (x) prediction is done by searching the entire training set for the most similar k instances. A distance calculation determines the most similar new input of k instance in the training dataset. Euclidean distance is the most popular distance calculation method for real world input variables. Euclidean distance is differences between a new point (m) and an existing point (m_i) across all input attributes j ,

$$(m, m_i) = \sqrt{\sum (m_j - m_{ij})^2} \tag{16}$$

Experimental setup and validation

The simulation setup and performances are described in this section. Simulation results show that the suggested approach can precisely classify exudates in retinal fundus images in terms of proliferative and non-proliferative.

Dataset description

DIARETDB1 is the diabetic retinopathy database which contains the images and corresponding ground truth. There is 89 color fundus images presented in the DIARETDB1 database and the 84 images are noted as mild signs (Ma) by experts. The 89 images are assigned as normal, moderate, mild, non-proliferative and severe proliferative which represents the progressive state of retinopathy. From the categories, the images are divided into training and testing as 28 and 61 respectively. The photos of the database DRIVE is collected from diabetic retinopathy program in Netherland. The people 25 to 90 years of age are participated in the screening program.

Quality measures

Some of the quality measures are evaluated to show the performance of the image. Figures 5, 6, 7 represents the exudate extraction, proliferative and non-proliferative type of hard exudates.

True negative rate (TNR)

True Negative Rate is the action of the proportion of negative events and it is also known as specificity. For a best case, it should be closer to one.

$$Specificity = \frac{TN}{TN + FN} \tag{17}$$

False positive rate (FPR)

The fraction of non-relevant features which remains recovered throughout of all non-relevant features. For a best case, it should be closer to zero.

$$FPR = \frac{FP}{FP + TN} \tag{18}$$

False negative rate (FNR)

FN means the positive instances happened for the detection case. For a best case, it should be closer to zero.

$$FNR = \frac{FN}{FN + TP} \tag{19}$$

Fig. 5 Extraction of exudates (a) input image (b) contrast enhancement image (c) optic disc detection (d) hard exudates

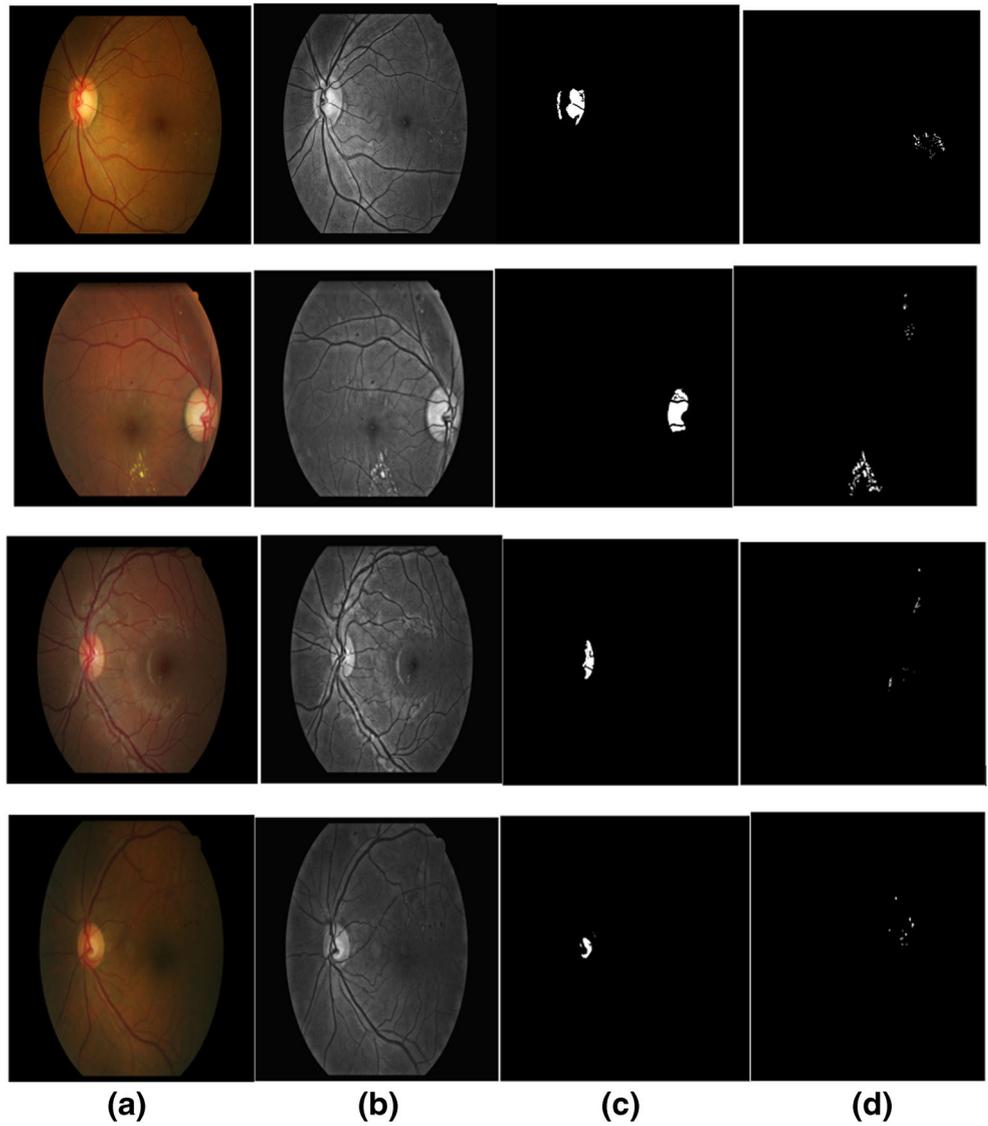


Fig. 6 Non-Proliferative diabetic retinopathy

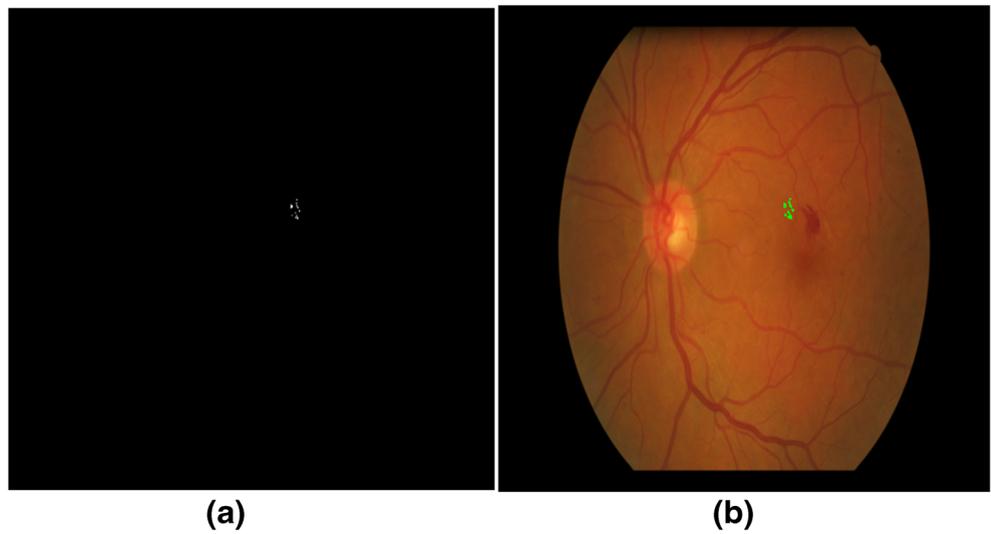
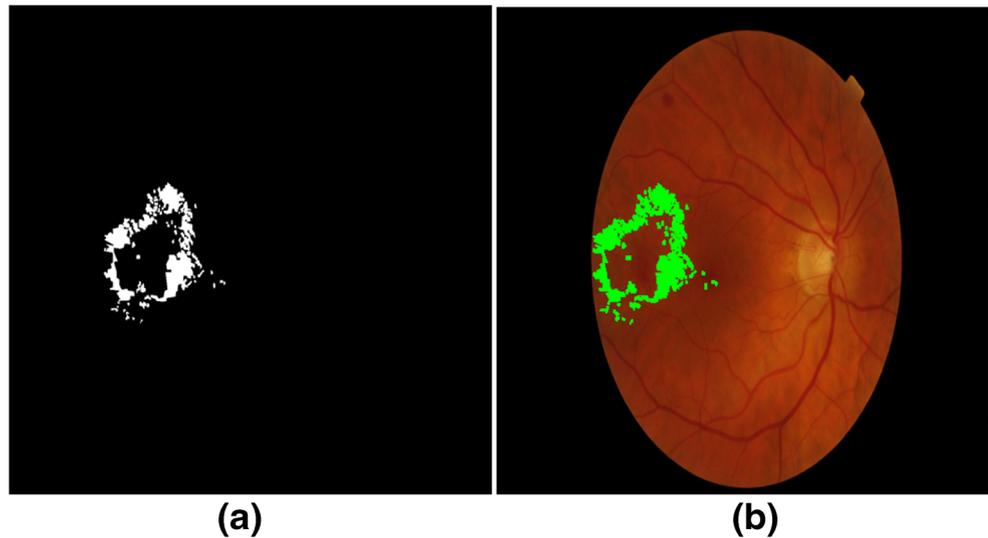


Fig. 7 Proliferative diabetic retinopathy



Precision

It determined the positive predictive value and maximum values of this precision which contains more information presented in the image.

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

Recall

It is the proportion of the measure of fitting information recovered to the aggregate entirety of suitable information.

$$Recall = \frac{TP}{TP + FN} \tag{21}$$

F₁-score

Precision and recall are combined in F1-score which is the mean average of both precision and recall. It can be expressed as

$$F_{1Score} = \frac{(1 + \beta^2) Recall * Precision}{\beta^2 (Recall + Precision)} \tag{22}$$

Accuracy

The proportion of the measure of suitable information recovered to the aggregate entirety of recovered information is known as accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{23}$$

Performance analysis

In the detection performance, the possible values are taken which got from one by one pixel classification computation. The total count of pixels presented in the test data are assessed initially. Then the evaluation is done for correct classification in the number of pixels in terms of ROC curve, sensitivity and

Table 1 Comparison table for proposed approach with the existing approaches

References	Extracted Features	Dataset	Accuracy (%)
Carla Pereira et al. [18]	Shape	HEI-MED	97%
T.jayaet al. [19]	Color, energy	Canon CR6-45NM	93%
Sundararajet al. [20]	Color, shape, size	DIARETB1	98.5%
Harry prattet al. [21]	Shape	Kaggle	75%
UsmaanAkramet al [22]	Shape, intensity, statistics	DRIVE, STARE, DIARETDB	97%
Proposed	Shape, intensity, energy, color, texture	DRIVE, DIARETDB1	99.34%

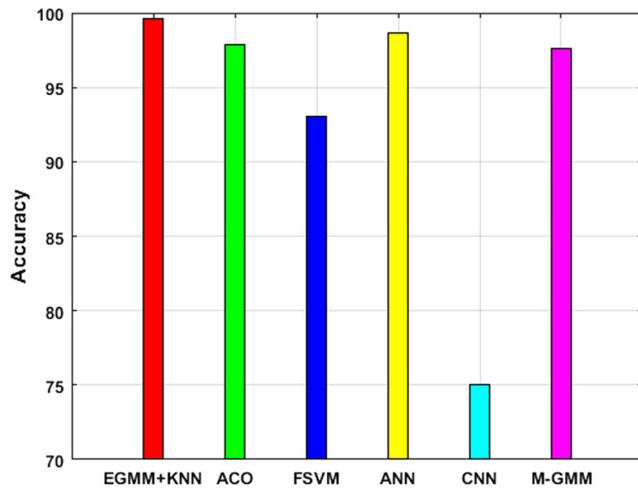


Fig. 8 Accuracy analysis

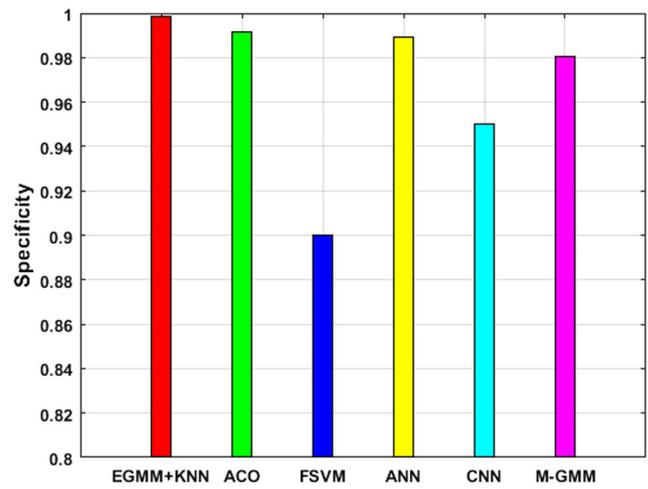


Fig. 10 Specificity analysis

specificity. Accuracy is the important performance which describe the overall detection system performance.

Various summary measures are described in the Table 1 which indicates the classifier performance. The KNN classifier provides an accuracy of 99.7253 which is better than the conventional approaches SVM, ANN and CNN classifiers. The proposed system has delivered a sensitivity of 99.42 and specificity of 99.87. KNN classifier has minimized misclassification rate of 0.02. The way toward recognizing hard exudates is troublesome and testing errand on the grounds that, the recently shaped vein is irrelevantly little in size and not unmistakably obvious. In test examination, acquiring an exact presumed locale in retinal picture is a significant issue on the grounds that, amid experimentation, the surface and the shading highlights got from the first picture expels the variances caused by veins, fluctuating picture quality and light. The following performance analysis corresponds to various classifiers shows KNN is higher in terms of accuracy, sensitivity, specificity and misclassification rate. The analysis

shows that the KNN algorithm will be a global solution for hard exudates detection with increased the detection accuracy and reduced misclassification rate. The above Table 1 compares the proposed method with five other existing approaches. The proposed approach is robust than the previous approaches by achieving 99% accuracy with DRIVE and DIARETDB1 database.

The accuracy is calculated for the proposed method using KNN classifier and it is compared with the existing methods. Figure 8 shows that the EGMM+KNN classifier has the highest accuracy of 99.34% by the combined features. Figures 9 and 10 illustrates the analysis of sensitivity and specificity which are compared with five relevant existing concept. It shows better performance for EGMM+KNN than the existing approaches such as ACO, FSVM, ANN, CNN and M-GM. The misclassification rate analysis is represented in Fig. 11 which indicates EGMM+KNN provides minimum misclassification rate. The ROC curves for evaluation on a per image basis are depicted in Fig. 12. For the per image basis the

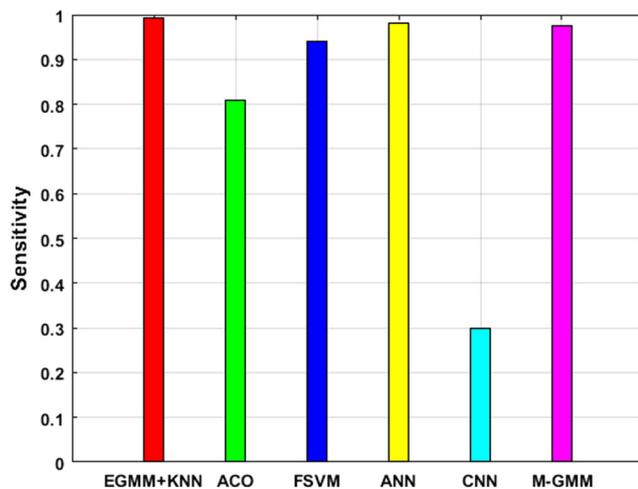


Fig. 9 Sensitivity analysis

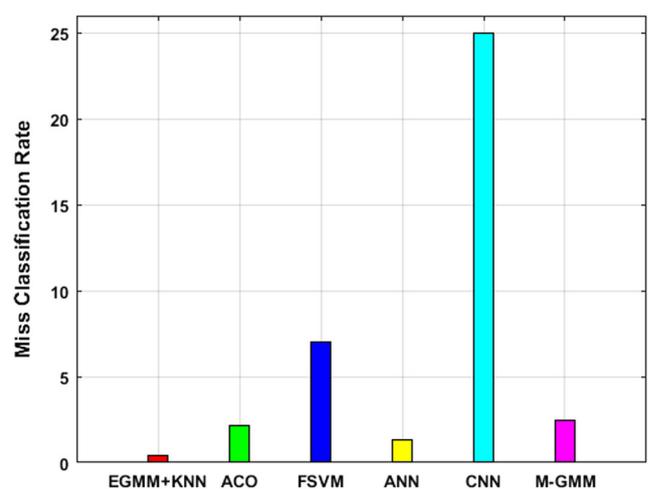


Fig. 11 Misclassification rate analysis

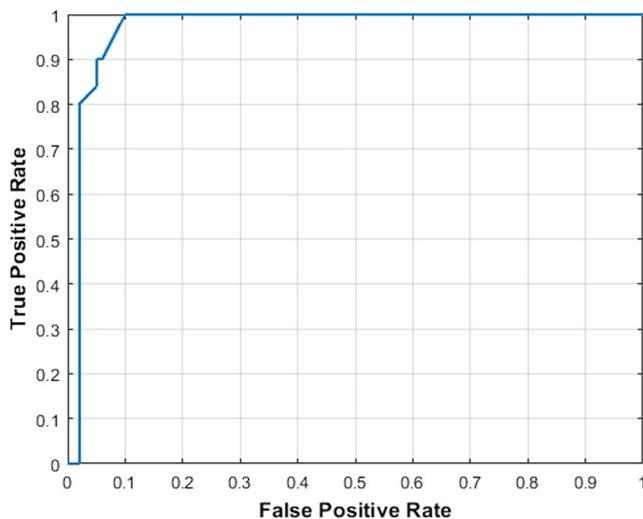


Fig. 12 ROC curve on image basis

Area Under Curve value is 0.9863 with maximum accuracy gives a sensitivity of 1.00 and a specificity of 0.94.

Conclusion and future works

In this research, a prediction approach is proposed which predicts the early signs of diabetic retinopathy termed as hard exudates. The supervised learning method starts the training process to predict the hard exudate. Then it detects the hard exudates through testing the retinal image. The discrimination between exudates and non-exudate pixel is done with the feature sets such as colour and texture. Initially, optic disc segmentation is done with edge detection and morphological process. Then the important features are selected by applying bio-inspired optimization algorithm. Then an accurate classifier KNN used to predict the hard exudates. From that the hard exudate is classified as proliferative and non-proliferative. The proposed approach provides better performance while compare the approach with state-of-art classifiers. The sensitivity achieved by proposed method is 0.891 in DIARETDB1 dataset. Our outcomes exhibit that the framework is appropriate to supplement the DR screening and might be used to help the ophthalmologists in their everyday practice. In future, the feature selection will be improved with better optimization approaches and classifier performance will be updated using deep learning approaches to enhance the classification accuracy and minimize the classification error rate.

Compliance with Ethical Standards

Conflict of Interest This paper is communicated only this esteemed journal with the knowledge of co-author.

Ethical Approval This article uses only standard databases not with any human participants or animals.

References

1. Haloi, M., Dandapat, S. and Sinha, R., A Gaussian scale space approach for exudates detection, classification and severity prediction. arXiv preprint arXiv:1505.00737, 2015.
2. Raj, A.M and Mani, S.A., Retinal Abnormality Risk Prediction Model: A hybrid approach based on vessel characteristics and exudates. Artificial intelligence and evolutionary computations in engineering systems, Springer, New Delhi, 803–818, 2016.
3. Ajwahir, M. I.S., Rajamani, K, and Sadhar, S.I., A novel technique for splat generation and patch level prediction in diabetic retinopathy. Annual conference on medical image understanding and analysis, Springer, Cham 50–59, 2017.
4. Marin, D., Gegundez-Arias, M. E., Ponte, B., Alvarez, F., Garrido, J., Ortega, C., Vasallo, M. J., and Bravo, J. M., An exudate detection method for diagnosis risk of diabetic macular edema in retinal images using feature-based and supervised classification. *Med. Biol. Eng. Comput.* 56(8):1379–1390, 2018.
5. Li, G., Zheng, S. and Li, X., Exudate detection in fundus images via convolutional neural network. International forum on digital TV and wireless multimedia communications, Springer, Singapore 193–202, 2017.
6. Zabihollahy, F., Lochbihler, A. and Ukwatta, E., Deep learning based approach for fully automated detection and segmentation of hard exudate from retinal images. In Medical Imaging 2019: Biomedical applications in molecular, structural, and functional imaging, International Society for Optics and Photonics, 10953: 1095308, 2019.
7. Abbasi-Sureshjani, S., Dashtbozorg, B., Romeny, B.M.H. and Fleuret, F., Boosted exudate segmentation in retinal images using residual nets. Fetal, infant and ophthalmic medical image analysis, Springer, Cham 210–218, 2017.
8. Otálora, S., Perdomo, O., González, F. and Müller, H., Training deep convolutional neural networks with active learning for exudate classification in eye fundus images. In Intravascular imaging and computer assisted stenting, and large-scale annotation of biomedical data and expert label synthesis, Springer, Cham 146–154, 2017.
9. Liu, Q., Zou, B., Chen, J., Ke, W., Yue, K., Chen, Z., and Zhao, G., A location-to-segmentation strategy for automatic exudate segmentation in colour retinal fundus images. *Comput. Med. Imag. Graph.* 55:78–86, 2017.
10. Manoj Kumar, S.B., Manjunath, R. and Sheshadri, H.S., Feature extraction from the fundus images for the diagnosis of diabetic retinopathy. Emerging research in electronics. Computer Science and Technology (ICERECT), 2015 International Conference on IEEE, 240–245, 2015.
11. Deshmukh, A. V, Patil T. G., Patankar, S. S. and Kulkarni, J. V., Features based classification of hard exudates in retinal images. Advances in computing, Communications and Informatics (ICACCI), 2015 International Conference on IEEE, 1652–1655, 2015.
12. Agurto, C., Murray, V., Yu, H., Wigdahl, I J., Pattichis, M., Nemeth, S., Barriga, E. S., and Soliz, P., A multiscale optimization approach to detect exudates in the macula. *IEEE J. Biomed. Health Inform.* 18(4):1328–1336, 2014.
13. Dashtbozorg, B., Mendonça, A. M., and Campilho, A., Optic disc segmentation using the sliding band filter. *Comput. Biol. Med.* 56: 1–12, 2015.
14. Akram, U. M., Tariq, A., Khan, S. A., and Javed, M. Y., Automated detection of exudates and macula for grading of diabetic macular edema. *Comput. Methods Prog. Biomed.* 114(2):141–152, 2014.
15. Patil, P., Shettar, P., Narayankar, P. and Patil, M., An efficient method of detecting exudates in diabetic retinopathy: Using texture edge features. In advances in computing, Communications and

- Informatics (ICACCI), 2016 International Conference on IEEE 1188–1191, 2016.
16. Balakrishnan, U., Venkatachalapathy, K., and Marimuthu, G. S., A hybrid PSO-DEFS based feature selection for the identification of diabetic retinopathy. *Curr. Diab. Rev.* 11(3):182–190, 2015.
 17. Karthikeyan, R., and Alli, P., Feature selection and parameters optimization of support vector machines based on hybrid glowworm swarm optimization for classification of diabetic retinopathy. *J. Med. Syst.* 42(10):195, 2018.
 18. Pereira, C., Gonçalves, L., and Ferreira, M., Exudate segmentation in fundus images using an ant colony optimization approach. *Inform. Sci.* 296:14–24, 2015.
 19. Jaya, T., Dheeba, J., and Singh, N. A., Detection of hard exudates in colour fundus images using fuzzy support vector machine-based expert system. *J. Digit. Imag.* 28(6):761–768, 2015.
 20. Franklin, W. S., and Rajan, S. E., Diagnosis of diabetic retinopathy by employing image processing technique to detect exudates in retinal images. *IET Image Process.* 8(10):601–609, 2014.
 21. Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., and Zheng, Y., Convolutional neural networks for diabetic retinopathy. *Proc. Comput. Sci.* 90:200–205, 2016.
 22. Akram, U. M., Khalid, S., Tariq, A., Khan, S. A., and Azam, F., Detection and classification of retinal lesions for grading of diabetic retinopathy. *Comput. Biol. Med.* 45:161–171, 2014.
 23. Mirjalili, S., Saremi, S., Mirjalili, S. M., and LDS, C., Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Syst. Applic.* 47:106–119, 2016.
 24. Chen, L.-F., Liao, H. Y. M., Ko, M.-T., Lin, J.-C., and Yu, G.-J., A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recogn.* 33(10):1713–1726, 2000.

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