

Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy

Romany F. Mansour¹

Received: 3 July 2017 / Revised: 11 August 2017 / Accepted: 24 August 2017 / Published online: 31 August 2017
© Korean Society of Medical and Biological Engineering and Springer-Verlag GmbH Germany 2017

Abstract The high-pace rise in advanced computing and imaging systems has given rise to a new research dimension called computer-aided diagnosis (CAD) system for various biomedical purposes. CAD-based diabetic retinopathy (DR) can be of paramount significance to enable early disease detection and diagnosis decision. Considering the robustness of deep neural networks (DNNs) to solve highly intricate classification problems, in this paper, AlexNet DNN, which functions on the basis of convolutional neural network (CNN), has been applied to enable an optimal DR CAD solution. The DR model applies a multilevel optimization measure that incorporates pre-processing, adaptive-learning-based Gaussian mixture model (GMM)-based concept region segmentation, connected component-analysis-based region of interest (ROI) localization, AlexNet DNN-based highly dimensional feature extraction, principle component analysis (PCA)- and linear discriminant analysis (LDA)-based feature selection, and support-vector-machine-based classification to ensure optimal five-class DR classification. The simulation results with standard KAGGLE fundus datasets reveal that the proposed AlexNet DNN-based DR exhibits a better performance with LDA feature selection, where it exhibits a DR classification accuracy of 97.93% with FC7 features, whereas with PCA, it shows 95.26% accuracy. Comparative analysis with spatial invariant feature transform (SIFT) technique (accuracy—94.40%) based DR feature extraction also confirms that AlexNet DNN-based DR outperforms SIFT-based DR.

Keywords Computer-aided diagnosis · Diabetic retinopathy · Deep neural network · AlexNet DNN · Convolutional neural network · Gaussian mixture model · Linear discriminant analysis · SVM

1 Introduction

Science and technology can be stated as a “blessing for human society” only when it makes human life more secure, healthier, livable, and comfortable. The immense anticipations for robust CAD systems have been witnessed globally. To meet global healthcare demands, CAD systems have emerged as vital technologies to assist earlier disease detection and diagnosis decision. Diabetes has become one of the most common and widespread health issues globally that occurs when the key functional organ, pancreas, malfunctions by secreting insufficient insulin, gradually influencing the human retina. Recently, the World Health Organization (WHO) has released a report which predicts that diabetes would become the 7th major death-causing health issue by 2030. The malfunctions caused by diabetes result into a disease called diabetic retinopathy (DR). Diabetic retinopathy has emerged exponentially globally, which can be annotated by futile glucose metabolism that increases the probability of long-term infection, especially in human retina. DR is one of the most generic and common diabetic eye issues or diseases, which emerges because of the damages in the retinal blood vessels. In major cases, blood vessels swell and leak fluid. In some cases, it blocks blood vessels completely, and as a result, certain abnormal blood vessels emerge over the retina during proliferation, which eventually causes loss of vision. Primarily, these malfunctions result in complete blindness and visual impairments. The vital information on

✉ Romany F. Mansour
romanyf@aun.edu.eg

¹ Department of Mathematics, Faculty of Science, New Valley
- Assiut University, Assiut, Egypt

retinal pathological transformation can be identified by means of ocular fundus images. Exudates are the predominant and generic symptoms of DR. Typically, micro-aneurysms and hemorrhages take place in the human retina, and exudates are secreted on it. The shape, size, and overall appearance of these features signify the level of disease.

Proliferative diabetic retinopathy (PDR) represents a phase of retinopathy where blood vessels grow. On the other hand, non-proliferative diabetic retinopathies (NPDR) refer to the initial phase of the DR and are categorized into three prime phases, namely, mild, moderate, and severe NPDR. Generally, mild NPDR is characterized by the existence of minimum micro-aneurysm, moderate indicates the existence or occurrence of hemorrhages, micro-aneurysms, and hard exudates, and the severe phase refers to the complete impasse of the retinal blood vessels. Thus, DR is identified by identifying key features such as, exudates, hemorrhages, and micro-aneurysms in human retina. Some of the dominant features of the DR and their respective shape and sizes are illustrated in Fig. 1.

The types of diabetic retinopathy compliant with its severity of DR are presented in Fig. 2.

The early detection of DR can play a vital role in assuring successful diagnosis and effective treatment. To characterize or classify the severity of the disease, DR employs weighting of various key features and their respective locations [1]. No doubt, it is an extremely time-consuming task for traditional practitioners or the clinicians. Developing a novel CAD system for DR can be of paramount significance for efficient DR diagnosis. However, the dominating intricacy in the process is to extract significant features such as exudates that typically resemble optic disk and possess similar color and size. Similarly, micro-aneurysms also look similar to the retinal blood vessels, especially in terms of color and propinquity with retinal blood vessels. The sophisticated computing mechanism has enabled retrieving swift feature classifications once trained, thus facilitating the CAD

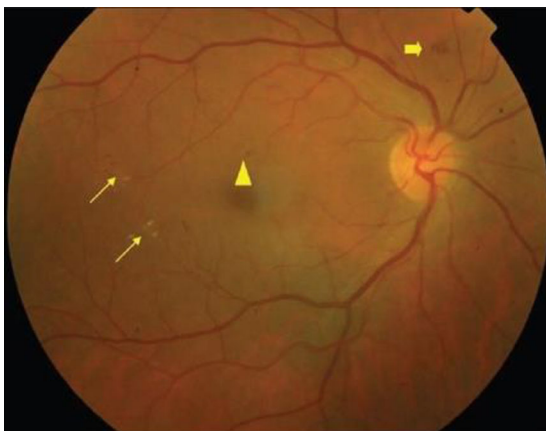


Fig. 1 NPDR: *thin arrows*—hard exudates; *thick arrow*—blot intra-retinal hemorrhage; *triangle*—micro-aneurysm

solution to the practitioners for better diagnosis decision. The efficacy of CAD solutions for DR characterization and severity labeling has been an open research domain in computer imaging systems [2, 3]. Recently, some efforts have been made to detect the features of DR using machine learning schemes such as support vector machines (SVM) and k-NN classifiers [4]. The development of convolutional neural networks (CNNs) has provided an attractive research domain in biomedical imaging system, particularly for image analysis and interpretation. Approaches were also developed decades before (1970s) to enable better imaging system and its analysis [5] to assist applications and effectiveness of approaches to solve certain task [6]. However, the breakthrough was observed only after the emergence of neural networks, where the dropouts were implemented to enable better outcome [7] and rectified linear units and graphical processor units (GPUs) were used to enable more effective computing systems, to solve image recognition purposes [8]. Recent studies [9, 10] have stated that the large CNNs can be applied to deal with highly complex image recognition applications with multiple object classes and can be effectively employed in major state-of-the-art image classification purposes such as the Image Net and COCO issues [9, 10]. However, enhancing the CNN to deal with complex problems such as DR has always been the motivation for researchers. With this motivation, in this paper, a CNN-based deep neural network (DNN) model has been developed for a DR CAD system. Unlike existing works, we proposed a multilevel optimization measure which, at first, tries to identify precise regions of interest (ROIs) and respective detection to enable optimal feature extraction using AlexNet DNN. Furthermore, to reduce computational overheads, different feature selection and dimensional reduction schemes have been applied. To examine the efficacy of the proposed DR CAD solution, a five-class classification (normal, mild NPDR, severe PDR, moderate NPDR, and PDR) has been performed using a SVM classifier.

The other sections of the presented manuscript are divided as follows. Section 2 presents related work, Sect. 3 discusses the contributions and proposed method, and Sect. 4 discusses the implementation/system model and associated processes. Section 5 discusses the results obtained. In Sect. 6, the overall research conclusions are presented. References used in this work are given at the last part of the paper.

2 Literature survey

Considering the significance of blood vessel segmentation for efficient DR, the authors in [11] applied different morphology and segmentation techniques to detect blood vessels, hard exudates, and micro-aneurysms. To extract

Fig. 2 Fundus images signifying different levels of diabetic retinopathy



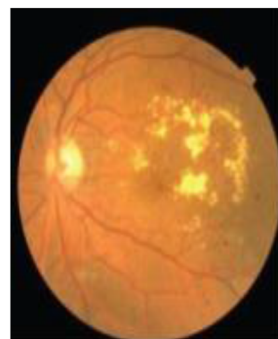
Without DR



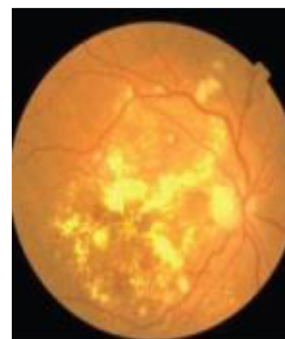
Early Diabetic Retinopathy



Mild NPDR



Moderate NPDR



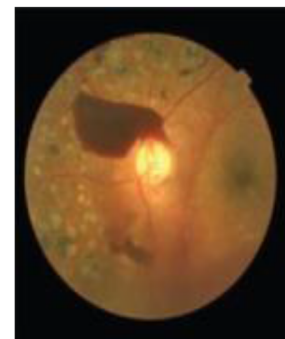
Severe NPDR



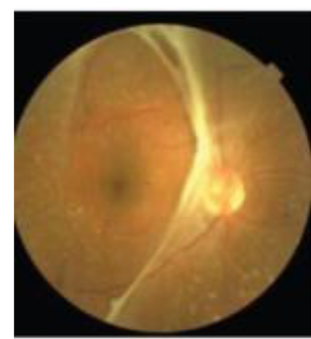
PDR and Neurovascularization



PDR with Vitreous hemorrhage



PDR with Vitreous hemorrhage and PLM



Vitreoretinal traction bands

features, they [11] used Haar wavelet transform, which was followed by PCA-based feature selection. The authors suggested a back-propagation-based neural network for two-class classifications. Similarly, in [12], multilayer perception neural network (MLPNN) was used to detect DR. An automatic exudate detection approach was developed in [13], where optic disk (OD) segmentation was performed using a graph cuts algorithm. The authors in [13] used the invariant moments H_u to extract the feature vector and NN-based two classes; exudate and non-exudate classification was

performed. Majority of these artificial neural networks (ANN)-based approaches [11–13] do not address the overfitting issues, particularly for large-scale fundus images. In [14], the authors applied optic disc identification for exudates and micro-aneurysm extraction-based DR, where they performed a five-class classification: mild, moderate, severe, NPDR, and PDR. To localize exudates, authors used a genetic algorithm [15]. To localize the exudates and other lesions in a fundus image, authors [16] applied the intersection of abnormal thickness in blood vessels. Similarly, in

[17], k-means clustering and fuzzy inference system (FIS) were applied for DR. In [18], emphasis was made on reviving an automatic CAD system for DR anomaly detection and its severity assessment. Authors [19] used eigenvalue analysis from Hessian matrix to detect micro-aneurysms. In [20], dynamic shape features of blood vessels were used for micro-aneurysm and hemorrhage detection. In [21], emphasis was made on detecting and classifying neovascularization originated owing to PDR. A technique called JSEG was used in [22] to detect exudates and abnormalities in a macular region. The authors in [23] performed the characterization of the parafoveal hemodynamics allied with DR along with the adaptive optics scanning laser ophthalmoscopy and computational fluid dynamics approach. Their prime objective was to identify an optimal feature to perform earlier DR prediction based on hemodynamics in blood vessels. In [24], the authors explored the efficacy of geometric features, especially related to the retinal vascular changes during a diabetic condition to predict DR. A matched filter with minimum cross entropy threshold (MCET) was applied to detect blood vessels [25] so as to enable an effective DR solution. They applied MCET for blood vessel detection in real time, enabling normal and abnormal retina classification. Authors [26] developed a hemorrhage detection scheme which, at first, detected blood vessels and eliminated so as to localize hemorrhage candidate. A generic CAD model with retinal micro-aneurysm and exudate detection, and an SVM and k-NN-classifier-based DR classification was used in [27]. Discrete-wavelet-transform-based feature extraction was performed in [28] to identify DR diseases in fundus images. In addition, they applied the PCA algorithm to reduce feature dimension, which was followed by Naïve Bayes-based classification. In [29], the Hurst exponent was used to estimate the fractal dimension (FD) that was applied for DR purposes. Morphology and texture analysis approach was used in [30] to detect DR features, like blood vessels, hard exudates, etc., in colored fundus images. To characterize DR severity level, authors used SVM-based feature classification [30, 31]. Texture features, along with exudates and micro-aneurysm features, were applied for DR classification [32, 33]. Authors used SVM to perform five-class DR level classification. The authors in [34] used ANN and pixel intensity values to perform DR classification, where the accuracy was 83.5%. The areas of exudates and blood vessels were used in [35] to perform DR characterization.

3 Our contribution

This research intends to develop an efficient automated computer-aided (CAD) system for DR. To enable a robust CAD solution for DR purposes, enhancing a major

comprising process level, such as data preparation, data quality enrichment, noise reduction, retinal blood vessel detection, ROI segmentation, feature extraction, and feature resizing and classification, is necessary. At first, pre-processing techniques such as noise removal and image resizing have been applied, which enriched fundus image quality and suitability for further processes. Once performing pre-processing, ROI segmentation has been performed, for which an adaptive-learning-rate-based enhanced GMM (E-GMM) algorithm has been developed. Here it should be noted that, to exhibit DR classification, different features, including hard exudates, blot intra-retinal hemorrhage, and micro-aneurysms, have been considered as target regions to be segmented so as to characterize the fundus image as DR or non-DR or other classes. Considering the diversity of features in this research, the KAGGLE DR dataset [36] has been applied. Now, realizing the fact that accurate and precise ROI identification and associated feature extraction can lead to accurate classification, the segmented region is processed for connected component analysis (CCA). The CCA method removes those image components that seem connected to the target region, which, however, are insignificant toward classification. For example, there can be pixel regions which may be connected to the DR traits or features, but may not give any significant characteristic to signify its role in DR classification. Thus, applying CCA and removing insignificant pixels, a precise ROI region is obtained, which is further processed for feature extraction using CNN. Unlike conventional feature extraction approaches such as wavelet transform and Gabor filter, in this work, the AlexNet DNN technique was applied to extract features from the CCA-processed segmented ROI. AlexNet DNN applied the CNN concept to extract highly dimensional features.

The proposed AlexNet model extracts multidimensional features at the fully connected (FC) DNN layers, such as FC6 and FC7, which is supposed to have more significant feature information. Considering the generic DNN approaches where the probability of over-fitting and degraded accuracy can be ignored, particularly for large-scale datasets, the proposed model applies CaffeNet DNN in conjunction with AlexNet that plays a vital role in feature extraction (for large-scale data) as well as execution with general-purpose computing systems. Practically, owing to higher unannotated data, performing DNN learning and classification is often a tedious task, and therefore, to alleviate such problem, AlexNet incorporates multilayered DNN, which is over large-scale labeled datasets to perform accurate DR classification. In the proposed model, the features are obtained at each layer of the trained DNN (convolutional layer-1 to layer-5 and FC 6 and FC7). The layered architecture of the DNN can be

found in Fig. 7. The higher layer of DNN model retrieves 4096-dimensional features that make it more informative to enable sufficient features for accurate DR classification. While exploring literatures, it has been found that spatial invariant feature transform (SIFT) technique can perform better than generic wavelet transform or Gabor filter techniques, and it has been applied to perform feature extraction over a CCA-processed segmented concept region. The discussion of SIFT-based feature extraction can be found in Sect. 4. As already stated, AlexNet DNN retrieves highly dimensional features at FC6 and FC7 that is significantly bulky in size, which could limit the computational efficiency of the proposed work. Therefore, two well-known dimensional reduction or feature selection methods, namely, principle component analysis (PCA) and linear discriminant analysis (LDA), have been applied. This process significantly assures optimal feature selection that eventually supports accurate classification. Once performing feature selection or dimensional reduction, the finally obtained features, also called feature vectors, have been projected to the polynomial kernel-based SVM classifier, where a tenfold cross-validation is applied to perform five-class classifications: no DR, mild DR, moderate DR, severe DR, and proliferative DR. The overall implementation flow can be found in Fig. 2.

Before discussing the implementation of the proposed DR method, a glance of the terms used and their nomenclatures is given as follows (Table 1):

4 System model

This section primarily discusses the overall proposed system and its implementation to achieve the optimal DR solution. Our proposed CAD solution for DR classification includes the following sequential steps. Figure 3 depicts the implementation model of the proposed DNN-based CAD system for DR.

- A. Data collection,
- B. Pre-processing,
- C. Concept region or ROI detection,
- D. Feature extraction,
- E. Dimensional reduction, and
- F. Diabetic retinopathy classification.

The discussion of the applied methodology is presented as follows.

4.1 Data preparation

Considering the objective of developing an automatic CAD system for DR, a standard DR image dataset named KAGGLE [36] has been used. KAGGLE benchmark data

Table 1 Nomenclatures

Nomenclature	
CAD	Computer-aided diagnosis
DR	Diabetic retinopathy
DNN	Deep neural network
ROI	Region of interest
SIFT	Spatial invariant feature transform
WHO	World Health Organization
PDR	Proliferative diabetic retinopathy
NPDR	Non-proliferative diabetic retinopathies
OD	Optic disk
CNNs	Convolution neural networks
GPU _s	Graphical processor units
SVM	Support vector machine
MLPNN	Multi layer perception neural network
FIS	Fuzzy inference system
FD	Fractal dimension
GMM	Gaussian mixture model
CCA	Connected component analysis
E-GMM	Enhanced GMM
PCA	Principle component analysis
LDA	Linear discriminant analysis
PDF	Probability density function
CONV-1	Convolutional layer-1
3D	Three-dimensional
PCS	Principle components
MEF	Most expressive features
MDF	Most discriminating features
LLE	Locally linear embedding
FV	Fisher vector
RBF	Radial basis function
ANN	Artificial neural network
FC	Fully connected
CONV	Convolutional (Layer)
w_k	Weight factor
x	Value of a pixel at an instant
$p(x)$	Gaussian mixture
α	Rate of learning
λ	Threshold
σ_k	Lowest standard deviation
β	Fixed sized learning rate
μ_k	Average density
$\gamma_k(t)$	Adaptive learning rate
F_V	Feature vector
$\sum_k = \sigma_k I$	covariance matrix
I_{ICW}	Intra-class scatter matrix
I_{IOS}	Inter-class scatter matrix
C	total classes
μ_i	Mean of the average vector of a class i
M_i	Total number of samples in class i

Table 1 continued

Nomenclature	
W	Eigenvectors of $I_{ICW}^{-1}I_{IOS}$
E	Level of error tolerance
x_i	Training vector
B	Background subtraction

provides a large set of high-resolution fundus images taken under varied imaging conditions. It contains a total of 35,126 fundus images, which are labeled as left or right, signifying whether the image belongs to the left eye of the patient or the right. Table 2 presents the fundus image distribution in the KAGGLE dataset.

These fundus images are taken from different cameras and under different conditions that may influence the visual appearances, thus demanding an optimal retinal blood vessel detection and ROI localization scheme. In this dataset, a fraction of the fundus images is presented in such manner that an individual can see the retina anatomically (i.e., macula on the left and optic nerve on the right for the

right eye). The other fundus images are shown as one can see by means of a microscopic condensing lens. It should be noted that the other type of fundus images presents the inverted image as usually seen in live eye test. In general, there are two approaches in assessing whether an image is inverted. These are the following:

1. The fundus image can be inverted when the macula (it presents the tiny dark central region) is little higher as compared to the optic nerve midline. In case the macula is lower than the optic nerve midline, it is not inverted.
2. In case of any notch sign (square, circle, or triangle) on the side of the fundus image, then it is not inverted. Otherwise, the fundus image is inverted.

Similar to other real-world dataset [37–39], the KAGGLE [36] DR fundus images also possess certain noise, where the image can be out of focus, over-exposed, or under-exposed. Considering huge data elements and associated noise complexities, in this paper, it is intended to develop a mechanism that could predict DR even in the presence of noise and disturbances. Before processing

Fig. 3 Proposed algorithmic implementation structure

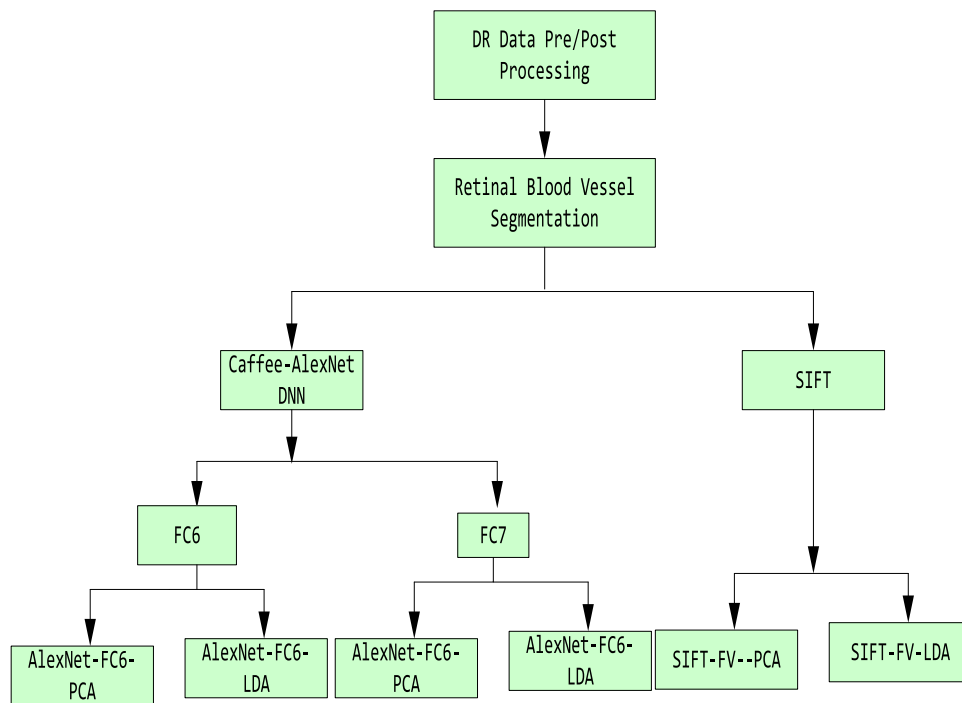


Table 2 KAGGLE benchmark data distribution

Class	DR classification	Fundus images	Percentage (%)	Imbalanced ratio
0	Normal	25,810	73.48	1.01
1	Mild NPDR	2443	6.96	1.84
2	Moderate NPDR	5292	15.07	1.26
3	Severe NPDR	873	2.48	2.76
4	PDR	708	2.01	2.89

blood vessel segmentation of ROI localization, the fundus images are resized to 512×512 , which is a valid input for AlexNet DNN-based feature extraction. In addition, it makes computation accurate and also reduces complexity. Figure 4 represents the resized data elements.

Retrieving the resized data, data pre-processing has been performed. A brief of the applied pre-processing phase is presented as follows.

4.2 Pre-processing

To enable optimal DR/non-DR characterization, being a sensitive data enabling a fundus image optimal for processing, image pre-processing is vital. The pre-processing approach refers to outstandingly vital same time finishing arrangement for micro-aneurysm vicinity in the pictures. The key features, like micro-aneurysms, would Verwoerd little and circular to nature. Completing pre-processing without suppressing whatever picture qualities should measure and figure out micro-aneurysms and other significant features accurately [40–42]. In this paper, the suggestion made by Antal [40] was applied to ensure optimal DR/non-DR-type characterization. Antal [40] applied an ensemble-based model to enhance the detection of micro-aneurysms (MAs) in the human eye. To perform MA

detection, both pre-processing and candidate region detection and extraction were applied. Considering the interchangeable nature of pre-processing methods, authors suggested to perform pre-processing before exhibiting candidate region detection and feature extraction. It avoids the variations in the original fundus image characteristics. To deal with noisy images, observing their suggestions and justifiable outcomes in [40], we have used histogram equalization that makes data suitable for further feature extraction and classification. More details of pre-processing can be found in [40]. Once performing pre-processing, detection of the blood vessels and associated key feature identification are necessary. In this paper, before processing feature extraction, at first, blood vessel extraction has been performed, which has been followed by associated ROI localization. To achieve this, blood vessel segmentation has been performed. A brief of the segmentation approach applied is presented in the following sub-section.

4.3 Concept region or ROI detection

This section primarily discusses the process of blood vessel segmentation applied to extract region of interest in fundus images. To perform blood vessel segmentation, in this paper, the background subtraction method has been

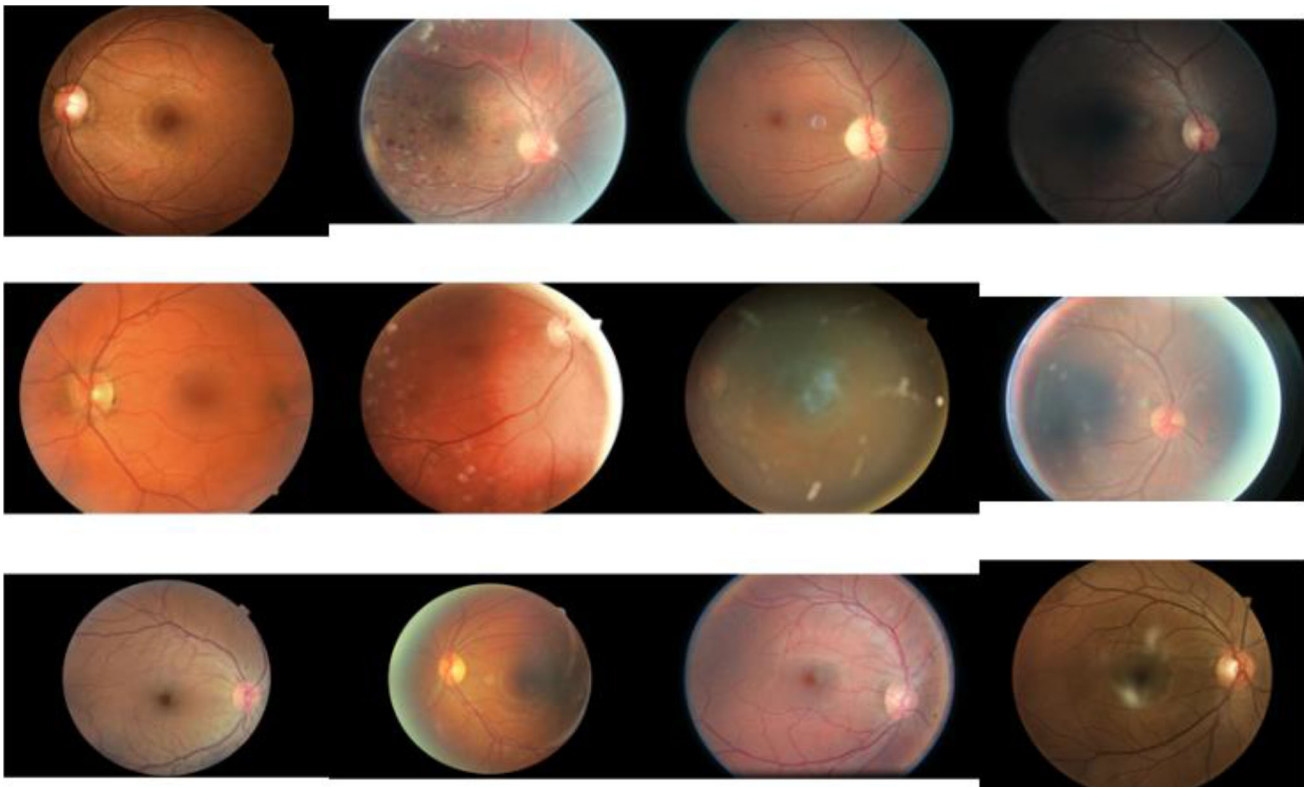


Fig. 4 Represents the resized data elements

applied. Here, it should be noted that the prime objective of retinal blood vessel segmentation is to distinguish other DR features such as hard exudates, MAs, and hemorrhage. In an array of background subtraction approaches, the Gaussian mixture model (GMM) algorithm has exhibited a better performance. Considering its effectiveness, unlike traditional algorithm, in this paper, an enhanced GMM algorithm has been developed, which has been applied to perform background subtraction and associated retinal blood vessel extraction for further DR/non-DR-type characterization of the fundus images. The discussion of the proposed GMM algorithm is discussed as follows:

4.3.1 Adaptive-learning-based GMM for background subtraction

Considering the complex color and textural features of the human retina, the traditional threshold-based GMM cannot be effective, and hence, to enable optimal blood vessel detection and associated feature extraction measure, in this paper, an enhanced adaptive-learning-rate-based GMM has been developed. In fact, GMM is a pixel-based approach that assures precise background subtraction. Let x be the value of a pixel at an instant. Then, to calculate the probability density function (PDF) of x , a GMM model can be used, where PDF may encompass the addition of all associated Gaussians. The PDF of Gaussian mixture $p(x)$ with K components can be calculated using (1).

$$p(x) = \sum_{k=1}^K w_k N(x; \mu_k, \sigma_k), \quad (1)$$

where w_k represents the weight factor, $N(x; \mu_k, \sigma_k)$ signifies the normalized density of the average μ_k , and $\sum_k = \sigma_k$. I represents the covariance matrix. This condition considers that the red, green, and blue pixel values in the fundus image are independent and possess similar variances. In contrast, it seems impractical, and hence, the distribution of recently observed values of the individual pixel in the image can be characterized in the form of a Gaussian mixture. Here, a new pixel value is signified by means of one of the dominating components of the mixture model and taken into consideration to update the model.

Authors [43] applied variables to estimate the image background, where, at first, these variables were initialized with zero. If there is any similarity, like $x - \mu_j/\sigma_j < \tau$, with $j \in [1, \dots, K]$ and $\tau (> 0)$ as a threshold, the functional parameters of the GMM model are updated as follows:

$$w_k(t) = (1 - \alpha)w_k(t-1) + \alpha M_k(t) \quad (2)$$

$$\mu_k(t) = (1 - \beta)\mu_k(t-1) + \beta x \quad (3)$$

$$\sigma_k^2(t) = (1 - \beta)\sigma_{k(t-1)}^2 + \beta(x - \mu_k(t))^2, \quad (4)$$

where

$$\begin{cases} M_k(t) = 1 & \text{with matching element } j \\ M_k(t) = 0 & \text{others.} \end{cases}$$

In case of non-matching elements, the element with the least w_k is re-initialized, and these are updated as follows:

$$w_k = w_0, \mu_k = \mu_0, \sigma_k = \sigma_0.$$

As depicted in above Eqs. (2)–(4), α states the rate of learning, and thus, parameter β is retrieved using Eq. (5).

$$\beta = \alpha N(x; \mu_k, \sigma_k). \quad (5)$$

Parameter w_k is normalized in such way that it converges toward 1 or increases to the unit value. Authors [43] performed the sorting of Gaussians w_k/σ_k in their decreasing order to derive the background model, where GMM uses a threshold λ in relation to the sums of the weights to achieve the set $\{1, \dots, B\}$, where B is estimated using (6).

$$B = \arg \min_{K_B} \left(\sum_{k=1}^{K_B \leq K} w_k > \lambda \right). \quad (6)$$

Authors [43] performed the sorting of Gaussians w_k/σ_k in their decreasing order to derive the background model. With a decrease in the variances, the Gaussian distribution gains increases, thus making it more evident to perform background identification. Once estimating these GMM parameters, it becomes feasible to sort from the matched mixture distribution toward optimal (say, most probable) background distribution. This is due to the fact that only the matched models' relative value would change. In this approach, the ordering is performed in such manner that the most probable background distributions exist on the top, whereas the less probable transient background distributions incline toward the bottom. In this manner, the initial B distributions are selected as the background model. Mathematically, B is estimated as (6), where λ represents the output of the smallest fraction of the data accounted for by the background. It takes the "optimal" distribution till a significant section λ of the recent data is accounted for. With a smaller value of λ , the background model is typically called as unimodal, and in this case, applying merely the most probable distribution can make computation more efficient. On the contrary, a higher λ refers to the multimodal distribution which is usually caused by a repetitive background motion. Interestingly, in our research model, we apply fundus images with static features, and hence, unimodal (i.e., lower λ) makes our approach computationally efficient. Gaussians with the highest w_k and lowest standard deviation σ_k signify the background region of the image, and thus, when

identifying the specific background region, the subtraction can be performed. In numerous GMM approaches, the mean factors (μ_k, σ_k) are generally updated using fixed-size learning rate β [43]. Considering the application-specific scenario of retinal blood vessel detection and segmentation and other DR feature extraction, these existing static-learning-rate-based methods could not ensure optimal DR-ROI identification and might result to their exclusion, thus impacting the overall DR fundus image classification accuracy. To alleviate this issue, authors altered [44], where β was used at the early learning phase, which enabled adaptation under varying surface conditions. There can be a certain surface condition where the pixel might neither be the retinal blood vessel or the DR-ROI features nor the background, but may be classified in either category. It may cause inaccurate blood vessel detection and other DR-ROI identification, hence influencing DR accuracy. Now, recall the statements [44] that increasing β may result into significantly high-rate pixel feature variations and can make ROI detection vulnerable. Exploring mathematically, it can be found that the square of the difference between the average and pixel values can introduce significant changes that eventually might lead to textural feature variations till GMM saturation over a complete pixel color range. Though the adaptive background mixture model [43] can enable target region detection and tracking precisely even under varying background conditions (for example vehicle detection in traffic), however, considering minute, fine-growing surface features and highly intricate region identification and distinction such as exudates, its efficacy is still unexplored. Realizing the need for an efficient learning model for background subtraction when there can be varying background conditions, authors [44] developed an enhanced Gaussian mixture learning model. However, during convergence (while performing background subtraction), it undergoes saturation, and hence, to alleviate this issue, we incorporated adaptive learning rate (ALR)-based GMM that not only detects minute and fine-growing target features but also ensures accurate ROI detection even under varying background features. In the GMM model, at first, β is decoupled with other averaging components μ_k and σ_k . Here, an ALR $\gamma_k(t)$ is introduced, which updates μ_k iteratively using a probability parameter $R_k = N(x; \mu_k, \sigma_k)$ signifying whether a pixel is a part of the k th Gaussian distribution or not. Mathematically,

$$\gamma_k(t) = \gamma_k(t - 1) + \frac{K + 1}{K} R_k - \frac{1}{K} \sum_{i=1}^K R_i. \tag{7}$$

Here in (7), K represents the number of distributions. In (7), R_k presents the probability parameter that states whether a pixel can be a part of the k th Gaussian distribution. In

this manner, the probability parameters and the particular Gaussian distribution $\gamma_k(t)$ can be obtained, which, as a result, can enable fast and precise Gaussian average update [43]. It may strengthen the system to cope up with the fast textural variations and illumination condition variation, hence enabling accurate DR-ROI localization. Now, replacing the proposed learning rate γ_k as β in (3), it can be observed that the self-controlling update of the variance might prohibit Gaussian saturation. In contrast, a fast learning rate can raise the issue of degeneracy, as discussed above. In this paper, to remove such limitation, a semi-parametric scheme has been proposed to calculate the variance so as to enable quasi-linear adaptation. It is effective for the condition where a very small variation from the average and a reduced response may cause a higher variation. To accomplish it, the proposed model employs a sigmoid function given in (8).

$$f_{a,b}(x, \mu_k) = a + \frac{b - a}{1 + e^{-SE(x, \mu_k)}}, \tag{8}$$

where $E(x, \mu_k) = (x - \mu_k)^T(x - \mu_k)$ and S signifies a sigmoid slope controller. Now, replacing (8) in (2), the variance update can be obtained as (9).

$$\sigma_k^2(t) = (1 - \rho)\sigma_{k(t-1)}^2 + \rho f_{a,b}(x, \mu_k(t - 1)), \tag{9}$$

where $\rho = 0.6$ and $f_{a,b}(x, \mu)_{\mathcal{R}^+}$ confines σ_k to a defined space $\mathcal{R} \in [\frac{a+b}{2}, b]$. Here, it should be noted that, unlike Eqs. (2)–(4), in Eq. (9), in spite of β , we have used a symbol ρ that differentiates these terms for different cases. Here, $\rho = 0.6$ is assigned on the basis of the experimental outputs where the two real positive variables a and b are selected in such way that \mathcal{R} spread over one k th of the pixel range. The value of ρ is estimated by manual computation while fulfilling the above mentioned conditions. In this way, introducing the proposed GMM model, background subtraction has been performed over each fundus image. Figure 5 presents the retinal blood vessel segmentation process. As depicted in Fig. 5, the first column represents the original fundus image from the KAGGLE dataset. The other two consecutive columns present the gray images with noise filtering, which is performed as per [40]. Column 4 presents the subtracted blood vessels and associated internal geometries. The last column presents the segmentation images adjusted by CCA.

4.4 Retinal blood vessel region localization

To enhance retinal blood vessel detection and optimal feature extraction, in this paper, the CCA method has been applied, which considers the region, size, and location of the key DR features such as hard exudates, MAs, and

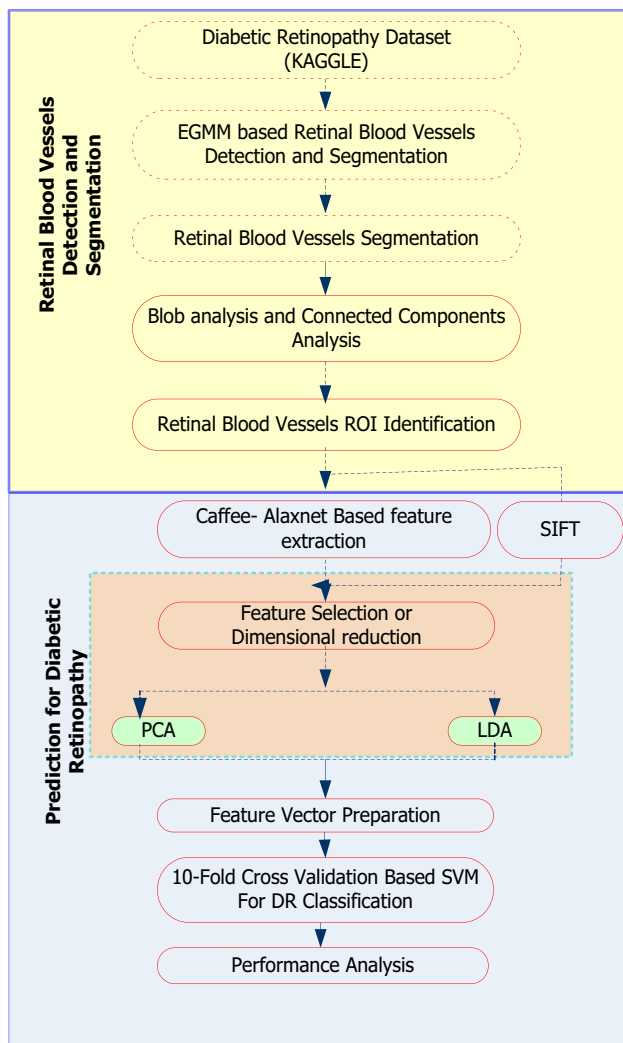


Fig. 5 Implementation model for the proposed diabetic retinopathy CAD system

hemorrhage, and their shape, size, and proximity to the normal retinal blood vessels. Here, a hypothesis that the disconnected region signifies the Gaussian components belonging to the ROI DR features has been taken into consideration. To use the key feature information, the dimension and proximity to the blood vessels have been normalized. The normalized dimensional features have been used as the width of the connected component region divided by the width of the vessels at the centroid of the connected region. Thus, applying this normalized width, it becomes flexible to compare the sizes of exudates, MAs, hemorrhage, etc., at distinct locations on the fundus image as well as to identify the abnormal retinal features (characterizing retinal disease). Thus, employing the enhanced GMM and CCA approaches, the exact blood vessel regions as well as other DR features have been localized, which has been followed by feature extraction.

4.5 Feature extraction

Once performing DR-ROI localization, its features have been extracted to perform automatic DR/non-DR-type characterization or classification. The DR/non-DR characterization and associated grading consist of recognizing very fine details, such as MAs, to some bigger features, such as exudates, on images of the eye. To enable an efficient CAD solution for DR with large-scale fundus image data, in this paper, a robust deep learning approach has been applied to extract significant DR-ROI features. Here, a well-known and robust image feature extraction model named AlexNet DNN, which functions based on the CNN principle, has been applied to extract key ROI features, including exudates and MAs, on the retina. LLE and other dimensionality reduction/embedding processes go beyond density modeling methods such as local PCA or mixtures of factor analyzers.

A brief of the AlexNet DNN is given as follows:

- AlexNet DNN

AlexNet is a multilayered DNN that functions based on the CNN concept and works on ImageNet. To explore the efficacy of a transferable deep neural network, in this study, the hypothesis stating that the higher layers of a DNN trained on a particular large-size labeled data can be generic enough for another image-based classification purpose [45] has been taken into consideration. Thus, in our DR image classification model, the features have been extracted for a specific layer inside a well-trained DNN model and are transferred further to perform the classification task. The applied AlexNet DNN model is trained on ImageNet [46]. Here, the AlexNet DNN architecture facilitates an optimal approach to learn rich midlevel fundus image features and corresponding semantic representations that signify DR perception. It enables our model to ensure high classification accuracy for targeted DR image classification. There is a state-of-the-art learning framework available [47]; however, considering ease of implementation and efficiency, in our model, we have applied CaffeNet [48]. It enables AlexNet functions on general-purpose computers. The brief of the AlexNet DNN mechanism is discussed as follows.

4.5.1 AlexNet DNN-based DR-ROI feature extraction

To perform DR classification, at first, feature extraction has been performed over the considered fundus images. Here, the AlexNet DNN-based model has been used to extract the features of the fundus images (standard KAGGLE fundus dataset). To enable a flexible implementation of the AlexNet DNN on a general-purpose computing platform, CaffeNet-assisted AlexNet DNN has been considered,

which has been trained over the localized DR-ROI features of the used DR dataset. In this work, to perform optimally significant feature extraction, a multilayered AlexNet DNN has been applied. Figure 6 presents the structure of the AlexNet DNN model used. Here, it can be seen that the applied DNN model contains five convolutional layers (CONV1–CONV5) and two distinct fully connected layers (FC6–FC7).

In this model, the initial low-level features can have generic features similar to the Gabor information or certain blob characteristics. On the other hand, the higher layers

contain a vital and more critical information that can play a significant role in DR classification. Considering these factors, in this paper, AlexNet, with its higher order features FC6 and FC7, has been used. Here, five convolutional layers (CONV1–CONV5) and two fully connected layers (FC6 and FC7) of AlexNet provide extracted DR-ROI features. The individual convolutional layer contains multiple kernels, which characterizes a three-dimensional (3D) filter connected to the results of the previous layer. Here, it can be seen that the fully connected layers contain multiple neurons encompassing the real positive value. Here, each

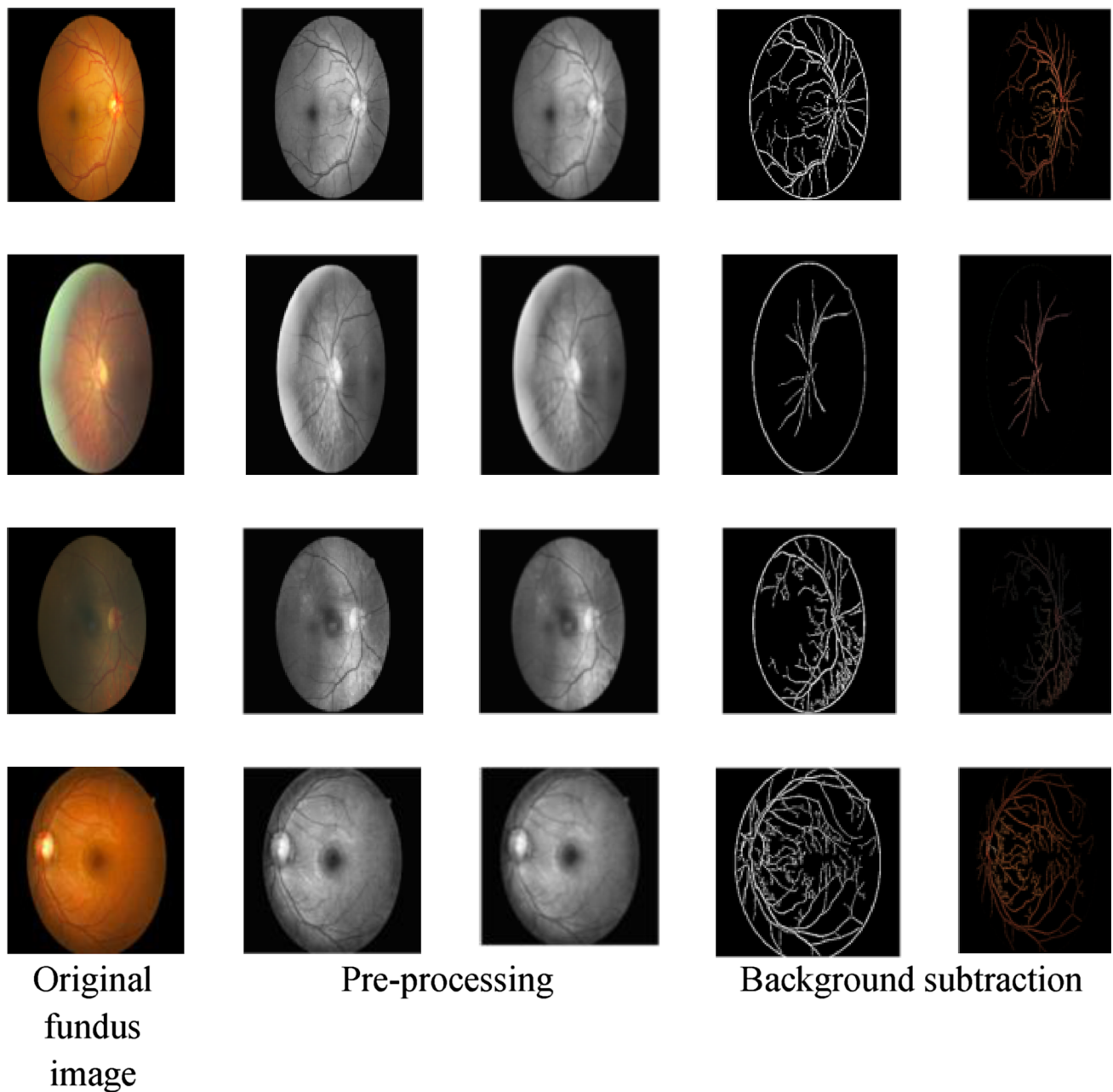


Fig. 6 Background subtraction for retinal blood vessels segmentation

individual neuron is connected to the entire neurons of the previous layer. Since higher layers of the CNN encompass critical information useful for feature classification, the FC6 and FC7 features have been applied. These high layer features facilitate 4096-dimensional ROI features that enable precise and accurate DR classification. The retrieved DR-ROI features are then converted into vector form, called feature vector $F_V = (f_1, f_2, f_3, \dots, f_{4096})$. The obtained feature vector F_V is then used for feature selection so as to enable dimensionally reduced data for efficient classification. Recently, authors have argued that SIFT can provide efficient ROI features to make accurate classification. Considering it as a reference technique, it has been applied to extract features from the detected ROIs on fundus images (Fig. 7).

4.5.2 SIFT-based DR-ROI feature extraction

In this paper, at first, we estimate the SIFT descriptors for each image, where each SIFT feature descriptor represents a 128-dimensional feature vector. The retrieved SIFT features have been processed for dimensional reduction using principle component analysis. Here, it should be noted that, in our model, we select only 50% of the 128-dimensional feature vectors, and thus, only 64 dimensions are reserved. It is then followed by Fisher encoding with the 32 Gaussian distributions that eventually generate a 4096-dimensional Fisher vector, which is equivalent to that generated by AlexNet DNN at the FC6/FC7 layers. Extracting 4096-dimensional features, it has been projected to the feature selection and classification to characterize fundus images as DR/non-DR type.

4.6 Feature selection and dimensional reduction

Extracting the features from AlexNet and SIFT descriptors, it has been projected for feature selection. Here, the prime intention is to use only significant features so that the computational overheads could be reduced. To perform feature selection or, in other words, dimensional reduction

in this paper, two algorithms, namely, PCA and LDA, have been applied individually.

4.6.1 Principle component analysis

Principle component analysis transfers a feature space of high dimension into lower dimension having the most significant features. It firmly rotates the axes of the p-dimensional space into a new position called principle axes in such manner that principle axes 1 possesses the maximum variance, axis 2 possesses the next highest variance, and so on. In extracted DR fundus image features, there can be a significant number of highly correlated feature elements, also called most expressive features (MEF). Therefore, the PCA feature selection process intends to transfer the feature elements to new feature vectors that are not correlated to each other. In this manner, PCA reduces a significant amount of feature elements having similar significance and selects MEF to form a new feature vector to be projected further for classification. It makes computation fast and accurate.

4.6.2 Linear discriminant analysis

Typically, the feature elements obtained from PCA are the MEF. In general, PCA applies MEF, whereas linear discriminant analysis uses the most discriminating features (MDF) for classification. Here, the irony is that MEF cannot be MDF universally. Dissimilar to the PCA approach, LDA exhibits feature selection automatically as to enable optimal feature space to be used for classification. For feature selection or dimensional reduction, LDA, at first, executes PCA, where DR-ROI, irrespective of its class label, is projected onto a single PC. To perform dimensional reduction (or feature selection), two distinct metrics, namely, intra-class scatter matrix I_{ICW} and inter-class scatter matrix I_{IOS} , have been estimated. Mathematically,

$$I_{ICW} = \sum_{i=1}^C \sum_{j=1}^{M_i} (y_j - \mu_i)(y_j - \mu_i)^T \quad (10)$$

$$I_{IOS} = \sum_{i=1}^C (\mu_i - \mu)(\mu_i - \mu)^T, \quad (11)$$

where C presents the total number of classes, μ_i represents the mean of the average vector of class i , and M_i refers to the total number of samples in class i . Here, the mean of the average vector, μ_i , is calculated as

$$\mu = \frac{1}{C} \sum_{i=1}^C \mu_i, \quad (12)$$

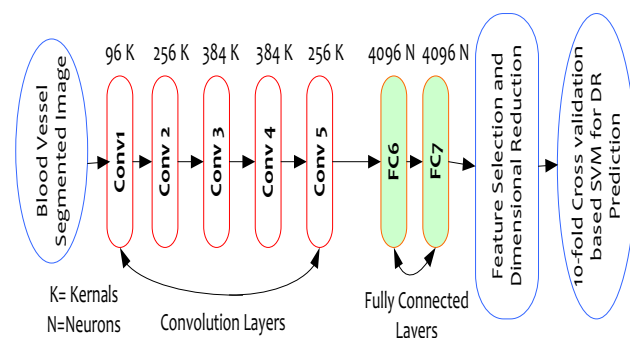


Fig. 7 AlexNet DNN architecture

where C presents the total number of classes. LDA emphasizes on increasing the inter-class scatter while minimizing the intra-class scatter. It is achieved by increasing a factor $\frac{\det|S_B|}{\det|S_w|}$. In case of a non-singular I_{IOS} matrix, this factor is increased when the column vectors of the projection matrix W are the eigenvectors of $I_{ICW}^{-1}I_{IOS}$. In this overall mechanism, W with $C - 1$ dimension allocates the training data onto a new feature space, called Fisher vector (FV). In other words, W is used to project all training samples on FV. Thus, the final FV to be used for classification is obtained as $F_{VR} = (f_{1R}, f_{2R}, f_{3R}, \dots, f_{4096R})$.

4.7 Classification

In this paper, radial basis function (RBF) kernel-based SVM has been used for training over extracted features. Here, the SVM algorithm has been trained over DR features (fundus image data) to obtain the maximum possible margin signifying the minimal value of w in Eq. (13).

$$\frac{1}{2}w^T w + E \sum \varepsilon_i, \quad (13)$$

where the parameters $\varepsilon_i \geq 0$ and E represent the level of error tolerance.

In this work, to exhibit two-class classification, the training feature vectors are grouped in certain labeled pairs $L_i(x_i, y_i)$, in which x_i represents the training vector. Here, the class label of training vector x_i is presented by $y_i \in \{-1, 1\}$. During the classification process, the hyper plane categorizes the maximum feasible points of the same class on the same side. In this paper, tenfold cross-validation is performed to ensure higher accuracy and reliability of the results. For DR classification, the test image data is processed for PCS retrieval, which is then processed for PCS classification using trained radial-basis-function-based SVM.

4.7.1 DR/non-DR-type characterization and classification

Final selected feature vectors (Table 2) have been projected and mapped for SVM-based classification. Considering non-linearity of the feature spaces, in our classification model, polynomial kernel-based SVM has been applied to perform five-class classifications. It classifies the fundus images into five classes, including normal, mild NPDR, severe PDR, moderate NPDR, and PDR. Furthermore, to achieve optimal classification accuracy, tenfold cross-validation has been performed. The retinal

blood vessels have been classified into two classes, namely, DR and non-DR.

5 Results and discussion

This section primarily discusses the results obtained in this paper. To examine the effectiveness of the DR CAD system, at first, the KAGGLE datasets have been collected. In this paper, a total of 35,126 images containing different fundus images with different levels of DR severity, features, trait location, orientation, etc., of exudates, MAs, hemorrhage, etc., were considered. To enable a computationally efficient process, at first, the fundus images were resized to 512×512 pixel sizes. Furthermore, to ensure optimal classification accuracy and overall performance, pre-processing was performed, which was followed by the blood vessel segmentation process. In this approach, unlike traditional GMM models, an enhanced learning-based GMM model in conjunction with CCA was executed. Thus, retrieving the ROI, the proposed CNN algorithm was applied to perform feature extraction. In this work, AlexNet DNN with CaffeNet model was used to perform feature extraction, where 4096-dimensional features at the higher layer (FC6 and FC7) were obtained. To enable a flexible AlexNet DNN implementation over general-purpose computers, CaffeNet has been used. The traditional ANN techniques may undergo over-fitting. This issue may be common, particularly with the applied data types where it may exhibit one-class classification, hence signifying that there is no sign of DR in the fundus image. The use of AlexNet can greatly alleviate this issue. To examine the effectiveness of the proposed AlexNet DNN-based feature extraction, a parallel feature extraction scheme, SIFT, was executed, which is similar to the AlexNet DNN extracted 4096-dimensional features. Initially, SIFT descriptors achieved 128-dimensional feature vectors. The extracted features were then projected onto PCA, where 50% of the initial PCs were applied (i.e., 64 dimensions) in conjunction with 32 Gaussian components distribution, and Fisher encoding generated 4096-dimensional features, equivalent to the AlexNet DNN features at the FC6 and FC7 layers.

To optimize computational efficiency and swift DR classification, extracted features were processed for feature selection and dimensional reduction using PCA and LDA algorithms. Once retrieving the optimal features, it was mapped to the feature vectors and was projected to the tenfold cross-validation-based SVM classifier, where a five-class classification, namely, DR, mild DR, moderate DR, severe DR, and proliferative DR characterization, was performed. To implement the overall model, VLFeat-0.9.20 toolbox was used, and all major executable functions were developed using MATLAB 2015a software tool.

Table 3 Accuracy of diabetic retinopathy two-class characterization

VI	Features	Classification accuracy (%)
FC6	ADR6-PCA	90.15
	ADR 6-LDA	97.93
FC7	ADR 7-PCA	95.26
	ADR -7LDA	97.28
SIFT	SIFT-FV-PCA	91.03
	SIFT-FV-LDA	94.40

* ADR signifies proposed GMM detection preceded AlexNet DNN

Since, in this proposed model, distinct training and testing data has been used, therefore, the tenfold cross-validation approach has been taken into consideration, which makes classification not only efficient but also reliable. Here, it should be noted that, since the proposed work intends to apply it on prepared data features for classification, therefore, performance assessment must be performed only with the prepared dataset. Considering diversity and robustness of the KAGGLE data, the suitability of the proposed system can be affirmed. However, as scope for further exploration, future authors can examine its suitability with other fundus datasets.

Tables 3 and 4 present the comparative performance assessment of the different DR approaches and respective DR image classification accuracy. In traditional approaches [29], authors applied image features such as contrast, correlation, energy, homogeneity, and entropy from the gray-level co-occurrence matrix (GLCM) of the fundus image. Authors applied these GLCM features for DR fundus image classification. No doubt, consideration of the inherent fundus image features makes this system exhibit a higher accuracy (94.60%); however, considering the complexities and closely related ambiguous features of the human eye often puts a question on its universality. As

depicted in Table 2, it can be seen that the proposed DR scheme outperforms majority of the existing systems and exhibits the highest classification accuracy (97.93%) with AlexNet FC6 layer's features with LDA selected feature vectors. Similarly, AlexNet FC7 features with LDA dimensional reduction (analogous, feature selection) also show a higher DR classification accuracy (97.28%) than PCA-based feature selection (95.26%). In the comparative performance assessment of Priya et al. in [49], where they used geometrical-feature-based ROI identification, followed by ANN-based feature extraction and SVM-based classification, to perform two-class DR classification, it can be found that they examined their performance only with a single fundus image. Similar to our proposed pre-processing technique, the authors [49] adopted grayscale conversion and adaptive histogram equalization. In addition, applying matched filter and fuzzy C-means clustering, they performed segmentation. Furthermore, to perform feature extraction, DWT was applied, and thus, eventual extracted features were classified using fuzzy C-means segmentation. From the pre-processed images, features were extracted for DR classification (using ANN). However, they could achieve an accuracy of merely 89.6%, which is much lower than the proposed method. No doubt, the efficiency of the DR classification is directly related to the optimal or accurate ROI identification, feature extraction, and its classification, and therefore, the lower accuracy (89.6%) of [49] can be caused by the inherent computational limitations of the fuzzy system, matched-filter-based segmentation that cannot ensure optimal ROI localization, especially for fundus images having a highly complex architecture and DR features. On the contrary, our proposed system intends to apply well-calibrated multilevel enhancement by first enriching image quality, ensuring optimal segmentation to lead the best possible DR-ROI feature extraction and classification. In practice, there can be huge data to be

Table 4 Comparative performance

Techniques	Classification accuracy (%)	Sensitivity	Specificity
FCM, shape, NN [50]	93.00	0.99	0.91
GLCM + SVM [27]	82.00	0.98	0.89
DWT + PCA [28]	95.00	0.99	0.88
HEDFD [29]	94.60	0.98	0.88
SVM + NN [49]	89.60	0.98	0.83
DNN [51]	96.00	1	0.98
ADR6-PCA*	90.15	0.98	0.87
ADR6-LDA*	97.93	1	0.93
ADR7-PCA*	95.26	0.96	0.93
ADR-7LDA*	97.28	1	0.99
SIFT-FV-PCA*	91.03	0.97	0.89
SIFT-FV-LDA*	94.40	0.94	0.93

* ADR states the proposed GMM-based AlexNet features

processed so as to enable accurate and swift DR classification. In addition, the impact of the limitations of the generic ANN algorithms such as local minima and convergence on feature extraction and its precision cannot be ignored. Therefore, the CNN-based DNN technique can be stated as the better alternative for the traditional NN algorithms. Considering time analysis, the simulation with MATLAB takes four hours to process the complete 35,126 images. On the contrary, GMM- and SIFT-based features have taken a relatively lower time (144 min); however, they perform at relatively compromised classification accuracy as compared to the AlexNet DNN-based DR model.

6 Conclusion

In the last few decades, the human society has witnessed numerous eye problems caused by diabetes globally. Conventional DR employ a manual fundus image analysis that needs experienced clinicians to identify, detect, and analyze the presence of even a minute or very small feature and its significance. In addition, it is highly time-consuming and difficult. To enable a better solution for DR, developing an automatic CAD solution can be vital; however, ensuring optimal accuracy has always been a question. In this paper, a CNN-based CAD system was developed for DR classification. The implementation of a multilevel optimization measure, including efficient pre-processing, enhanced GMM-based background subtraction, Caffee–AlexNet-based DNN for DR feature extraction, and *n*-fold cross-validation-based SVM for two-class classification, exhibited accurate DR classification. The implementation of the enhanced GMM in conjunction with CCA enabled efficient background subtraction and precise retinal blood vessel as well as key feature detection. As a result, this enabled the proposed system to perform a highly accurate DR classification. Unlike major existing approaches, where the complete fundus images are processed by CNN for classification, the proposed scheme extracted AlexNet features only for the identified or the identified ROI (blood vessels, hard exudates, micro-aneurysm, and hemorrhages). It significantly enhanced the efficiency of the CAD system. The implementation of AlexNet CNN with CaffeeNet enabled highly dimensional feature extraction that eventually provided a sufficient feature to characterize the fundus images into two classes, namely, DR and non-DR. The results revealed that the highly dimensional (4096-dimensions) features at the fully connected CNN layers (FC6 and FC7) contain a significant information to be used for DR classification purposes. This study also revealed that the selection of significant features can be vital to ensure a higher efficiency with minimum

time and processing resource requirements. It is of great significance for CAD-DR solutions. The proposed DNN model in conjunction with the polynomial SVM and ten-fold cross-validation exhibited the highest DR classification accuracy with AlexNet FC6 LDA features (97.93%). AlexNet FC7 features with LDA dimensional reduction (analogous, feature selection) also showed a higher DR classification accuracy (97.28%) than PCA-based feature selection (95.26%). AlexNet DNN outperformed other feature extraction approaches such as spatial invariant Fourier transform, which shows a DR classification accuracy of 91.03% with PCA feature selection and 94.40% with LDA feature selection. The overall results affirmed that AlexNet DNN with higher order CNN features can be an effective solution to enable efficient DR detection. In the future, multiclass classification can be performed to make the proposed CAD-DR system more effective and useful for diagnosis purposes. With a novel intention to explore the effectiveness of both the enhanced GMM and AlexNet DNN for DR applications, feature extraction was performed over GMM-based segmented and localized ROI. However, in the future, the comparison can be made with AlexNet features extracted from the original fundus image without applying GMM-based segmentation and ROI identification.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie House classification. ETDRS report number 10. *Ophthalmol.* 1991; 98:786–806.
2. Philip S, Fleming AD, Goatman KA, Fonseca S, Mcnamee P, Scotland GS. The efficacy of automated disease/no disease grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol.* 2007;91(11):1512–7.
3. Fleming AD, Philip S, Goatman KA, Prescott GJ, Sharp PF, Olson JA. The evidence for automated grading in diabetic retinopathy screening. *Curr Diabetes Rev.* 2011;7:246–52.
4. Mookiah MRK, Acharya UR, Chua CK, Lim CM, Ng E, Laude A. Computer-aided diagnosis of diabetic retinopathy: a review. *Comput Biol Med.* 2013;43(12):2136–55.
5. Fukushima K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern.* 1980;36(4):193–202.
6. Cun YL, Boser B, Denker JS, Howard RE, Hubbard W, Jackel LD. *Advances in neural information processing systems 2.* Citeseer. ISBN 1-55860-100-7; 1990. pp. 396–404.
7. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res.* 2014;15(1):1929–58.
8. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10).* 2010. pp. 807–814.

9. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. 2015; URL: [arXiv:1502.03167](https://arxiv.org/abs/1502.03167).
10. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2015; URL: [arXiv:1512.03385](https://arxiv.org/abs/1512.03385).
11. Prasad DK, Vibha L, Venugopal KR. Early detection of diabetic retinopathy from digital retinal fundus images. In: 2015 IEEE recent advances in intelligent computational systems (RAICS), Trivandrum; 2015. pp. 240–5.
12. Bhatkar AP, Kharat GU. Detection of diabetic retinopathy in retinal images using MLP classifier. In: 2015 IEEE international symposium on nanoelectronic and information systems, Indore; 2015. pp. 331–5.
13. Elbalaoui A, Boutaounte M, Faouzi H, Fakir M, Merbouha A. Segmentation and detection of diabetic retinopathy exudates. In: 2014 international conference on multimedia computing and systems (ICMCS), Marrakech; 2014. pp. 171–8.
14. Raman V, Then P, Sumari P. Proposed retinal abnormality detection and classification approach: computer aided detection for diabetic retinopathy by machine learning approaches. In: 2016 8th IEEE international conference on communication software and networks (ICCSN), Beijing, China; 2016. pp. 636–41.
15. Kaur A, Kaur P. An integrated approach for Diabetic Retinopathy exudate segmentation by using Genetic Algorithm and Switching Median Filter. In: 2016 international conference on image, vision and computing (ICIVC), Portsmouth; 2016. pp. 119–23.
16. ManojKumar SB, Manjunath R, Sheshadri HS. Feature extraction from the fundus images for the diagnosis of diabetic retinopathy. In: 2015 international conference on emerging research in electronics, computer science and technology (ICERECT), Mandya; 2015. pp. 240–5.
17. Jahiruzzaman M, Hossain ABMA. Detection and classification of diabetic retinopathy using K-means clustering and fuzzy logic. In: 2015 18th international conference on computer and information technology (ICCIT), Dhaka; 2015. pp. 534–8.
18. Wijesinghe A, Kodikara ND, Sandaruwan D. Autogenous diabetic retinopathy censor for ophthalmologists—AKSHI. In: 2016 IEEE international conference on control and robotics engineering (ICCRE), Singapore; 2016. pp. 1–10.
19. Sri RM, Rajesh V. Early detection of diabetic retinopathy from retinal fundus images using eigen value analysis. In: 2015 international conference on control, instrumentation, communication and computational technologies (ICCICCT), Kumaracoil; 2015. pp. 766–9.
20. Seoud L, Hurtut T, Chelbi J, Cheriet F, Langlois JMP. Red lesion detection using dynamic shape features for diabetic retinopathy screening. *IEEE Trans Med Imaging*. 2016;35(4):1116–26.
21. Mudigonda S, Oloumi F, Katta KM, Rangayyan RM. Fractal analysis of neovascularization due to diabetic retinopathy in retinal fundus images. In: E-health and bioengineering conference (EHB), 2015, Iasi; 2015. pp. 1–4.
22. Gandhi M, Dhanasekaran R. Investigation of severity of diabetic retinopathy by detecting exudates with respect to macula. In: 2015 international conference on communications and signal processing (ICCSP), Melmaruvathur; 2015. pp. 0724–9.
23. Bernabeu MO, Lu Y, Lammer J, Aiello LP, Coveney PV, Sun JK. Characterization of parafoveal hemodynamics associated with diabetic retinopathy with adaptive optics scanning laser ophthalmoscopy and computational fluid dynamics. In 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Milan; 2015. pp. 8070–3.
24. Leontidis G, Al-Diri B, Wigdahl J, Hunter A. Evaluation of geometric features as biomarkers of diabetic retinopathy for characterizing the retinal vascular changes during the progression of diabetes. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Milan; 2015. pp. 5255–9.
25. Kuri SK. Automatic diabetic retinopathy detection using OMFR with minimum cross entropy threshold. In: 2015 international conference on electrical engineering and information communication technology (ICEEICT), Dhaka; 2015. pp. 1–6.
26. Bharali P, Medhi JP, Nirmala SR. Detection of hemorrhages in diabetic retinopathy analysis using color fundus images. In: 2015 IEEE 2nd international conference on recent trends in information systems (ReTIS), Kolkata; 2015. pp. 237–42.
27. Lachure J, Deorankar AV, Lachure S, Gupta S, Jadhav R. Diabetic retinopathy using morphological operations and machine learning. In: 2015 IEEE international advance computing conference (IACC), Bangalore; 2015. pp. 617–22.
28. Alver S, Ay S, Tetik YE. A novel approach for the detection of diabetic retinopathy disease. In: 2015 23rd signal processing and communications applications conference (SIU), Malatya; 2015. pp. 1401–4.
29. Rao MA, Lamani D, Bhandarkar R, Manjunath TC. Automated detection of diabetic retinopathy through image feature extraction. In: 2014 international conference on advances in electronics, computers and communications (ICAIECC), Bangalore; 2014. pp. 1–6.
30. Du N, Li Y. Automated identification of diabetic retinopathy stages using support vector machine. In: 2013 32nd Chinese control conference (CCC), Xi'an; 2013. pp. 3882–6.
31. Gandhi M, Dhanasekaran R. Diagnosis of diabetic retinopathy using morphological process and SVM classifier. In: 2013 international conference on communications and signal processing (ICCSP), Melmaruvathur; 2013. pp. 873–7.
32. Adarsh P, Jeyakumari D. Multiclass SVM-based automated diagnosis of diabetic retinopathy. In: 2013 international conference on communications and signal processing (ICCSP), Melmaruvathur; 2013. pp. 206–10.
33. Adarsh P, Jeyakumari D. Multiclass svm-based automated diagnosis of diabetic retinopathy. In: 2013 international conference on communications and signal processing (ICCSP). IEEE; 2013. pp. 206–10.
34. Gardner G, Keating D, Williamson T, Elliott A. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. *Br J Ophthalmol*. 1996;80(11):940–4.
35. Nayak J, Bhat PS, Acharya R, Lim C, Kagathi M. Automated identification of diabetic retinopathy stages using digital fundus images. *J Med Syst*. 2008;32(2):107–15.
36. <https://www.kaggle.com/c/diabetic-retinopathy-detection/data>.
37. <http://www.it.lut.fi/project/imageret/diaretdb1/#DATA>.
38. <https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set>.
39. <http://www.rodrep.com/data-sets.html>.
40. Antal B, Hajdu A. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Trans Biomed Eng*. 2012;59(6):1720–26.
41. Zuiderveld K. Contrast limited adaptive histogram equalization. In: *Graphics gems IV*, Academic Press Professional, Inc., 1994. pp. 474–85.
42. Abdelazeem S. Microaneurysm detection using vessel removal and circular Hough transform. In: *Proceedings of the nineteenth National radio science conference*; 2002. pp. 421–6.
43. Stauffer C, Grimson WEL. Adaptive background mixture models for real-time tracking. In *Proceedings of IEEE conference on computer vision and pattern recognition*, 1999; vol. 2, pp. 246–52.
44. Lee DS. Effective Gaussian mixture learning for video background subtraction. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(5):827–32.

45. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks?. In: Advances in neural information processing systems; 2014. pp. 3320–8.
46. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems; 2012. pp. 1097–105.
47. Razavian AS, Azizpour H, Sullivan J, Carlsson S. Cnn features off-the-shelf: an astounding baseline for recognition. 2014. arXiv preprint [arXiv:1403.6382](https://arxiv.org/abs/1403.6382).
48. Jia Y et al. Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the ACM international conference on multimedia. ACM; 2014. pp. 675–8.
49. Priya R, Aruna P. SVM and neural network based diagnosis of diabetic retinopathy. *Int J Comput Appl.* 2012;41(1):6–12.
50. Singh N, Chandra R. Automated early detection of diabetic retinopathy using image analysis techniques. *Int J Comput Appl.* 2010;8:18–23.
51. Haloi M. Improved microaneurysm detection using deep neural networks. Cornell University Library; 2015.