



Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy

Sraddha Das ^a, Krity Kharbanda ^a, Suchetha M ^{a,*}, Rajiv Raman ^b, Edwin Dhas D ^a

^a Centre for Healthcare Advancement, Innovation & Research, Vellore Institute of Technology, Chennai, India

^b Shri Bhagwan Mahavir Vitreoretinal Services, Sankara Nethralaya Eye Hospital, Chennai, India



ARTICLE INFO

Keywords:

Diabetic retinopathy
Maximum principal curvature
Hessian matrix
Squeeze-excitation
Bottleneck
Convolutional neural network

ABSTRACT

Diabetic retinopathy is ophthalmological distress, diabetic patients suffer due to clots, lesions, or haemorrhage formation in the light-sensitive region of the retina. Blocking of vessels leads, due to the increase of blood sugar leads to the formation of new vessel growth, which gives rise to mesh-like structures. Assessing the branching retinal vasculature is an important aspect for ophthalmologists for efficient diagnosis. The fundus scans of the eye are first subjected to pre-processing, followed by segmentation. To extract the branching blood vessels, the technique of maximal principal curvature has been applied, which utilizes the maximum Eigenvalues of the Hessian matrix. Adaptive histogram equalization and the morphological opening, are performed post to that, to enhance and eliminate falsely segmented regions. The proliferation of optical nerves was observed much greater in diabetic or affected patients than in healthy ones. We have used a convolution neural network (CNN) to train the classifier for performing classification. The CNN, constructed for classification, comprises a combination of squeeze and excitation and bottleneck layers, one for each class, and a convolution and pooling layer architecture for classification between the two classes. For the performance evaluation of the proposed algorithm, we use the dataset DIARETDB1 (standard Diabetic Retinopathy Dataset) and the dataset provided by a medical institution, comprised of fundus scans of both affected and normal retinas. Experimental results show that the proposed algorithm provides improved results, when compared to traditional schemes. The model yielded an accuracy of 98.7 % and a precision of 97.2 % while evaluated on the DIARETDB1 dataset.

1. Introduction

People with Diabetes face a medical condition called Diabetic Retinopathy (DR), which is of two types, (a) Non-Proliferative Diabetic Retinopathy (NPDR), the milder form, (b) Proliferative Diabetic Retinopathy (PDR), the advanced form. The initial signs of the DR are the exudates, which indicate the milder condition, i.e. NPDR. In NPDR, the patient will only have blurry vision, but as the disease develops, the retina then grows new blood vessels that highly affect the vision. These abnormal blood vessels may leak or bleed easily, causing blood clots/blobs to emerge in the retina. Damage to the network of vessels that nourish the retina is mostly the key cause of DR. In the superior stages of the PDR, the blood vessels are entirely blocked, forming lesions in the blood vessels. The most visible lesions which appeared are the microaneurysms and hemorrhages. Micro-aneurysms are the first visual symptom of DR, which appear as small round-shaped red dots in the fundus. Currently, a trained ophthalmologist detects the DR by assessing

the fundus image manually. There is a need for an automated DR screening system that detects the condition quickly and accurately. Various methods have been worked on to build a system with high accuracies, such as a unique unsupervised method. For the detection and segmentation of the fovea in the retinal image, pixel-level exudate detection using deep-learning-based algorithms are used. In deep learning, CNN is a class of deep neural network which contains different layers of neurons. Each neuron of a layer is connected to all the neurons of the next layer. CNN has wide applications in the field of image classification.

In this paper, we have implemented a scheme that can be used to detect the DR. For diabetic patients, the unusual growth of blood vessels in the eye is the most prominent trait. Therefore, estimating the branching retinal vasculature is an important aspect for ophthalmologists, to perform an efficient diagnosis. The common classification algorithm involves the steps such as pre-processing, segmenting the fundus scans of the eye to have a clear view of the branches, followed by

* Corresponding author.

E-mail address: suchetha.m@vit.ac.in (S. M.).

a classification algorithm. Fundus scans are commonly used for the diagnosis of DR. Fig. 1 shows the sample fundus scan images where the blood vessels can be easily diagnosed. Different algorithms are used in segmentation that includes thresholding, edge-based segmentation, region-based segmentation, etc. After referring to different techniques for segmentation, we worked on the technique of Maximal principal curvature, which utilizes the maximum Eigenvalues of the Hessian matrix. Also, for the classification process, different algorithms are commonly used which include support vector machine, clustering-based approaches such as Fuzzy C means (FCM), and K-means. Several classification techniques also exist that include the methods of attention-aware, self-attention, ConvLSTM, and One-dimensional deep attention convolution network.

Our proposed method includes a novel deep learning architecture to classify segmented fundus images like 'Normal' and 'Abnormal', i.e., healthy and diabetic retinopathy affected. Several approaches utilize the CNN to learn features in the retinal scans; we have worked on segmented images. For this, a method of finding maximum principal curvatures is deployed, which extracts the blood vasculature effectively. Process of adaptive histogram equalization and the morphological opening is performed to enhance and eliminate falsely segmented regions. To improve the performance of the classifier, we constructed a two-way classification CNN model. One way for classification contains a combination of the squeeze, excitation, and bottleneck layers. The other one includes convolution and pooling layer architecture. The architecture of both modules is independent and distinct to perform robust feature extraction. A test image is taken as input into both networks, and only if the classification is unanimous, the result is accepted. Due to this, the model gives a significantly better performance than prior methods.

The paper is structured in the following manner. Section II contains a description of related works in DR detection and classification, Section III elaborates on the working of the proposed algorithm. Section IV contains the results and discussion of the proposed work and finally, Section V concludes the paper.

2. Related works

Detection of eye-related diseases from fundus images is an emerging area for researchers these days. Especially, automated detection of DR has become an important topic for researchers working in the field of medical image processing and computer vision.

Several papers focus on the segmentation process because the classification result mainly depends on the segmentation output. Wrong segmentation results can lead to the wrong classification because the features are extracted from non-DR pixels. The following paper focuses on different segmentation approaches. Especially, the authors, Santosh et al. [1] presented a method to segment the retinal fundus images with remarkable accuracy. It proceeds in two stages: Pre-processing of the

input image and post-processing of the pre-processed image by using maximum principal curvature. Gehad et al. [2] proposed a blood vessel segmentation approach. The technique can be used in retinal image analysis to extract the retinal image vessels. Mathematical morphology and K-means clustering are used to segment the vessels. Smoothing operation on the retinal image using mathematical morphology to enhance the blood vessels and suppress the background information. Finally, the K-means clustering algorithm is used to enhance the image. The proposed approach is tested on the DRIVE dataset and is compared with alternative approaches. This algorithm succeeded with an average accuracy of 95.10 %. Orlando et al. [3] present a discriminatively trained method based on extended description and evaluation for blood vessel segmentation in fundus images. The problem statement for this paper is to overcome the difficulty when Standard segmentation priors such as a Potts model or total variation deal with thin and elongated structures. The model is trained in a way that parameters of the method are learned automatically through a structured output support vector machine. Also, the authors R.Manjula et al. [4] developed a paper that aims to employ image processing techniques to enhance and measure the dimensions of the retinal blood vessels. Three techniques have been implemented to carry out the process of segmentation; Gaussian method, mathematical morphology method, and multi-scale analysis method. Gaussian method uses a Gaussian resolution hierarchy with an objective to detect thick and thin vessels. It is a faster technique but it is suitable only for detecting thick vessels. The mathematical morphology technique is more suitable to detect the fine details of thin vessels with further precision. The third technique, the multi-scale analysis method, is preferable for invariant analysis along with the transformation of images as it detects the thick and thin vessels without noise. Memari et al. [5] propose an automatic retinal vessel segmentation that utilizes fuzzy c-means clustering and level sets. Contrast limited adaptive histogram equalization is used to contrast enhance the retinal images, while the noise is condensed through mathematical morphology techniques along with subsequently matching filtering steps that use Gabor and Frangi filters, which improve the blood vessel network preceding the clustering. For extracting an initial blood vessel network, a genetic algorithm enhanced spatial fuzzy c-means scheme is used. The segmentation is further refined by an integrated level set approach. This method of segmentation achieved a mean accuracy of 96.1 %. Also, Budai et al. [6] aim to reduce the running time of the algorithm and segmenting the vessels in the fundus image. A method is proposed to reduce calculation time, that achieves high accuracy and increase sensitivity against the Frangi method. The authors have also work to avoid potential problems like specular reflexes of thick vessels while building this approach. They used two public databases, DRIVE and STARE, each obtaining accuracy of 95.72 % and 93.86 % respectively.

The segmentation methods such as superpixels based segmentation, watershed segmentation, and active contour methods are also very

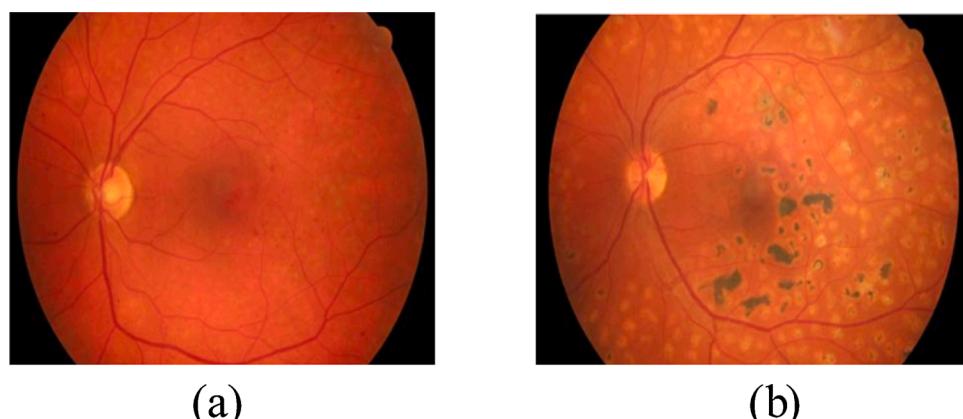


Fig. 1. Scans of fundus images (a) without DR (b) with DR.

popular segmentation schemes. In superpixels, based segmentation [7, 8], the input image is segmented into small homogeneous regions or superpixels. Superpixel segmentation provides an over-segmentation of a picture by gathering pixels into homogeneous bunches dependent on intensity, texture, and different other highlighted features. Every superpixel belonging to the same region label and arrangement of superpixels provides a map, which takes into account the local and nonlinear deformations of interesting regions. A small number of superpixels with more pixels reduce the computational cost and memory usage and make the superpixel-based feature more reliable and can minimize the risk of assigning wrong labels to the superpixels. The watershed transform [9,10] is a popular and reliable unsupervised model, applied to solve diverse image-segmentation problems, which is used by the field of mathematical morphology. This has been applied to solve various difficult problems of image segmentation, successfully. On a topographic surface, watershed transformation helps to determine the watershed lines. However, applying the watershed transform directly to the original image leads to over-segmentation. As the morphological watershed transformations are sensitive to noise and contrast in an image. The over-segmentation can be sufficiently serious to render incorrect segmentation results. Very often Active contour method (ACM) is used in digital images for several applications, including edge detection, segmentation, stereo matching, and shape recognition. It is considered as a most valuable tool for segmenting regions of interest (ROIs) in medical images. ACM [11,12] has two models, edge-based models, and region-based models. Edge-based models use local edge information to fit the boundaries of the approximated shape. Region-based active contour methods find the optimal energy for which the model fits the image best based on statistics calculated from subregions.

Several papers exist for detecting and classifying the DR present in the fundus images. For example, Dulanji et al. [13] exhibit that different morphological operations could be incorporated for accurate detection of exudates while focusing on optic disc detection. As it is the essential factor for the removal of false positives. Otherwise, the optic disc could be falsely recognized as exudates. They achieved sensitivity and specificity of 94.5 % and 88.46 % respectively. The authors Renoh et al. [14], introduces a unique unsupervised method to detect and segment OD and fovea from the retinal images. The proposed method consists of three steps, namely, coarse ONH center detection, fine-tuned ONH center and boundary detection, and fovea detection. They have proposed a method that automatically detects the optic disc (OD) using histogram-based template matching along with the maximum sum of vessel information in the retinal image. The optical disc detection accuracy came out to be 95 % and fovea detection accuracy to be 97.26 %. Also, Ramasubramanian et al. [15] work on hemorrhages and propose a solution for the efficient detection of hemorrhages without segmenting or eliminating any regions. Here a method for the automatic detection of hemorrhages in color retinal images is proposed and validated. The color retinal images used as an input that are captured from the diabetic patients are enhanced, using a bag of features based on intensity, color and texture, by an effective pre-processor. Finally, the features are classified with the help of a partial least square classifier. They obtained the AUC-ROC value of 0.98. The authors, Avula et al. [16] deploy a deep learning model to detect the exudates in the retinal images. Here they have tried to generate an image in which hard exudates are detected by predicting the class of every pixel. The model presented in this paper is developed using Tensorflow deep learning framework. For this method sensitivity and specificity attained was 98.29 % and 41.35 % respectively. Shuang et al. [17] focus on potential exudate candidate points that are first extracted with morphological ultimate opening methods and then the candidate points are passed for classification to the trained CNN deep networks. Before putting the input to the designed algorithm they first pre-process the image to remove the optic disc and blood vessels. The CNN network is trained on the 64*64 patches. They achieved a pixel-wise accuracy to be 91.92 %.

Shailesh et al. [18] proposes an improved blood vessel detection technique by morphological iterative process and develops an automated Optic Disc algorithm. They have worked to bring improved techniques for microaneurysm as well as hemorrhages detection, eventually contributing to the overall improvement in the early detection of DR. The method consists of five stages- pre-processing, detection of blood vessels, segmentation of optic disc, localization of fovea, feature extraction and classification. Pre-processing and blood vessel detection are performed using Mathematical morphology operation. Optic disc segmentation is done using Watershed and the classification of the diseases is done by a Radial basis function neural network. This paper developed sensitivity and specificity of 87 % and 93 % respectively. N. Yalçın et al. [19] proposed an approach of two steps. Step 1 is to perform pre-treatments to remove retinal images from different datasets and standardize them to size. In Step 2, classification was made by Convolutional Neural Network. Instead of creating the feature set manually as in other traditional methods, the deep learning network automatically constructs itself in a very short time by using the CPU and GPU in the training phase where 98.5 % success was attained. Shah et al. [20] came up with an AI-based algorithm, which helps to detect DR. The authors have validated using an internal dataset consisting of 1533 macula-centered fundus images and an external MESSIDOR dataset. Interobserver agreement using kappa value was calculated for both the sets and two out of three agreements for DR grading was considered as ground truth to compare with AI results. The outcome was a sensitivity of 90.4 % and specificity of 91 %. Xu et al. [21] presented a complex architecture containing a compound of stacked layers; 15 layers out of which 13 are convolutional layers and two fully connected layers. It is a two-class classification model, with each class having a precision value of 0.81 and 0.88 respectively. Also, the authors discovered that when pre-processing is done using contrast limited adaptive histogram equalization along with ensuring dataset fidelity by expert verification of class labels enhances the recognition of refined features. Andonova et al. [22] used the MESSIDOR database and proposed a CNN with 4 convolutional layers. Input images are first pre-processed, transformed, and normalized. This enhanced their quality. For classification purposes, they are split into multiple groups (clusters). This method developed max accuracy of 82.5 %. Also, Prentasic et al. [23] developed an algorithm based on CNN to detect exudates from the color fundus images. CNN or deep neural networks helps to segment neuronal membranes in electron microscopy. The convolutional neural network in this paper calculates the probability of a pixel in one of the two classes; exudate or non-exudate class. This method attains a sensitivity of 77 %. Adem K [24] proposed a hybrid approach of circular Hough transform and CNN algorithms for detecting the exudates. Three DiaretDB0, DiaretDB1, and DrimDB public datasets were used to assess. While testing with the images not included in the training set, it was found to have a correct classification ratio of 99.17 % in DiaretDB0, 98.53 % in DiaretDB1, and 99.18 % in DrimDB.

The following papers focus on different approaches that classify the DR using Neural Networks. The authors, Kejie et al. [25] put forward various concepts such as a shortcut connection and an attention mechanism for multi-task deep learning models which can be used to optimize the process of representation sharing. This Attention-aware Multi-task Convolutional Neural Network can automatically learn appropriate sharing via end-to-end training. To suppress redundant contents contained in the representations the attention mechanism is introduced in this paper. M. Li et al. [26] presents a discussion on the design of a cost-effective universal retinal fundus camera and the building of a new algorithm for the identification of prominent vision-threatening diseases. They also introduced an autoencoder neural network and a novel 3D self-attention convolutional neural network. Hemanth et al. [27] in their study included image processing with histogram equalization, and the contrast limited adaptive histogram equalization techniques and performed the classification using a convolutional neural network. 400 retinal fundus images from the MESSIDOR database were used to

validate the method which achieves a precision of 94 %. Shuyuan et al. [28] proposed an algorithm to automatically extract and classify the different features of the signals. In the first step, one-dimensional sparse filters are designed to learn hierarchical features of raw signals. Second, an attention layer is constructed to weigh and assemble feature maps, to derive more context-relevant representation. Wen-Shuai et al. [29] use long short-term memory (LSTM), as a special deep learning structure to improve the classification performance. Zhao et al. [30] have used visualized feature maps of DenseNet and CondenseNet CNN architectures. The authors also propose a relative-squeezing bottleneck design to improve the computing efficiency of Convolutional Neural Networks. Three benchmark datasets: CIFAR-10, CIFAR-100, and ImageNet are used to evaluate the method. Jun et al. [31] proposed a new algorithm called a deep residual squeeze and excitation network, to enhance the representation ability of the network, as a building blocks in DRSEN. Also, to reduce the parameters of the network, work has been done to improve the up-sampling module and the global residual pathway in the network. Lahmri [32] suggests SENet architecture, that generalizes extremely challenging datasets. In this process, deep learning convolutional neural networks are used for automatic feature extraction. And the Student *t*-test is applied to the high dimensional features set, which is extracted by CNN, to select the best ten features. Finally, to perform the classification task, the selected CNN-based features are fed to a nonlinear support vector machine tuned by Bayes optimization.

Instead of using a single-stage classifier as discussed in most of the classifiers, the proposed method includes robust two independent CNN modules for classification, containing the squeeze, excitation, and bottleneck units. The maximum principal curvature algorithm performs a highly efficient segmentation of the fundus images to extract the proliferative blood vasculature as the problem of vessel trench detection is categorized by high magnitudes of the curvature. The adaptive histogram equalization and the morphological opening processes are used to eliminate the wrongly segmented regions. In the early advanced stages of DR, detection of new blood vasculature growth serves as a preventive step before the blood vessels leak and hamper vision later on. Certain constraints entail as well. As stated in most of the articles about diabetic retinopathy, in the advanced conditions of the DR, the patients develop an abnormal growth of blood vessels, which tend to leak and hamper vision. Sometimes when patients are near the advanced stage of the DR or in milder conditions may not have enough proliferation of blood vessels to be detected.

3. Proposed methodology

Our proposed architecture involves two modules used for classification- the memory module and a central CNN. In the memory module, a squeeze operation squeezes a feature. Max-pooling blocks are a major part of the squeeze process that enhances the informational features and suppresses the inefficient features. Batch normalization layers expedite the convergence and thus improve the stability. Non-linear ReLU activation function models each, ‘Normal’ and ‘Abnormal’ data. A bottleneck layer of convolutional layers decreases the model’s complexity. The purpose of the bottleneck layer is to reduce the number of channels by a certain degree using a poor 1×1 convolution so that the latter 3×3 convolution has lesser parameters. The network at the end is widened again with another 1×1 convolution process where our bottleneck includes three convolution layers. The squeeze, excitation, and bottleneck units in the memory module perform robust feature extraction and reduce the overall complexity of the model. Powerful feature extraction is highly essential when dealing with such medical images that contain fine structures. It also enhances the CNN performance with only a minimal increase in the number of total parameters. Finally, ReLU processing is performed on the added results using an FC, followed by the classification layer after softmax activation. The other module, the central CNN, has a basic convolutional layer architecture ending in an FC layer. Two different network architectures are responsible for the

classification of the same image and finally categorizing a particular image to a certain category, only if both the modules give the same result. The images are segmented using the method of maximum principal curvatures, which efficiently extracts the fine blood vasculature. Then adaptive histogram equalization and the morphological opening are performed to enhance and discard incorrect regions. The various layers in the Memory module and bottleneck layer of the suggested deep-learning architecture are given in Fig. 2.

The retinal area in diabetic patients contains eminent vasculature. Our proposed algorithm for the detection and classification of DR in blood vessels caused due to diabetes has been explained below. Fig. 3 depicts the block diagram of the proposed DR detection and classification algorithm.

The proposed work has three stages for DR classification. The first stage is pre-processing, the second stage is segmentation and the third stage is classification. The algorithm of the proposed classification process can be summarized as shown below,

3.1. Algorithm

Step (i): Each image undergoes pre-processing before being used for any step ahead. Pre-processing involves resizing the image to a common size and converting it to a grayscale image format. Usually, the greyscale image is estimated from the R, G, and B values using the relation,

$$G = [0.299R + 0.587G + 0.114B] \quad (1)$$

Since we aim to extract the blood vessels and red clots; we provide higher weightage to red color when compared to other two-color components. Therefore, the greyscale conversion can be modified as,

$$G = [0.587R + 0.299G + 0.114B] \quad (2)$$

Step (ii): The second step is to segment the image; Eigenvalues of a Hessian matrix are calculated to calculate maximum principal curvature. The image that is acquired by scan machines can induce noise in the fundus images. To suppress such noise, a Gaussian filter using a Gaussian function is used to smoothen the image. The morphological opening is used to remove any gaps between pixels in the image.

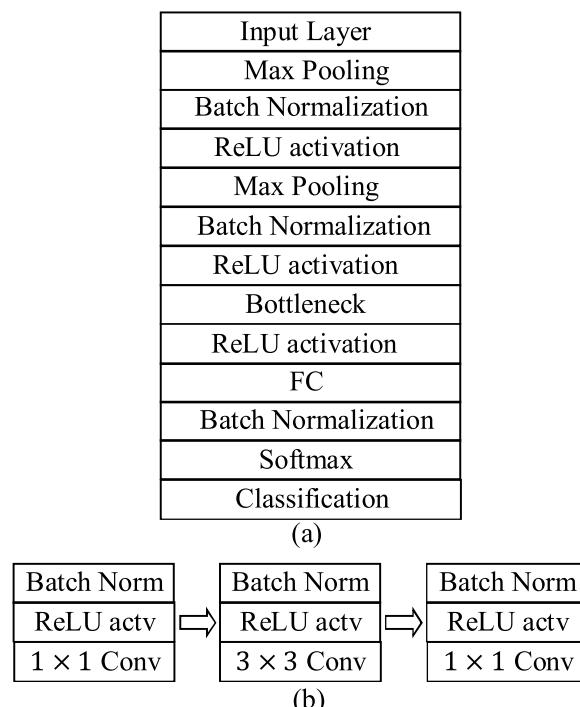


Fig. 2. Architecture of (a) Memory module and (b) Bottleneck layer.

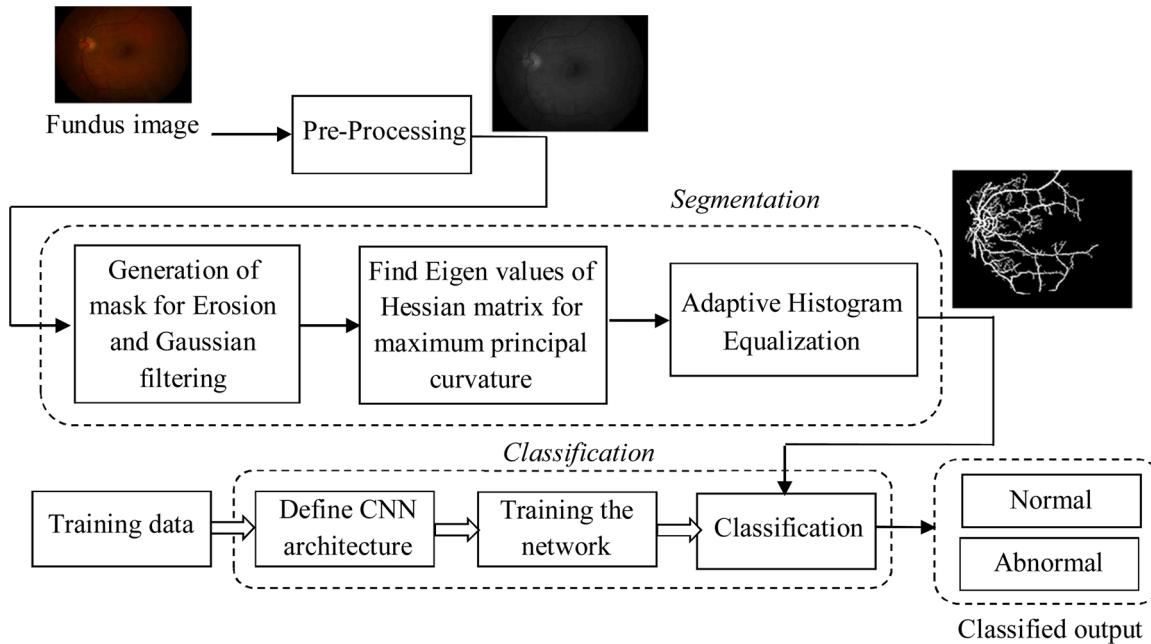


Fig. 3. Overview of the proposed algorithm.

Morphological operator erosion is used to remove the pixels that lie on the boundary which appear like small objects so that only substantive objects remain. An octagonal structuring element of size 24 is used which provides the optimal result. The blood vessels are detected so that further analysis can be done to detect the DR using the maximum principal curvature algorithm. The differences in the gradient of intensities about the neighborhood of a pixel best describe the principal curvature in which the vasculature with a ridge construct in the image emerges. A Hessian matrix can estimate the maximum principal curvature to the direction of the eigenvector which is described for every pixel value and can be defined as the highest Eigenvalue of the calculated second-order derivative. For a Hessian H , given as Eq. (4),

$$H = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix} \quad (3)$$

$Z(x,y)$ is a function of intensity values of the image. The second-order derivative of $Z(x,y)$ with respect to x is Z_{xx} and with respect to y is Z_{yy} . Differentiation of Z_x with respect to y results in mixed partial derivative Z_{xy} . For a unit vector $(a, b)^T$ in the y direction, its second-order form is calculated as shown in Eq. (5).

$$[A \ B] = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \quad (4)$$

Product of Eigenvalues of a Hessian produces its determinant. The eigenvector having the largest Eigenvalue gives the direction of maximum curvature. Eigenvalues or principal curvatures determine the extreme bending of a regular plane at every point or variations in the gradient of intensity with regards to the neighborhood of a pixel in which the upcoming vasculature is ridge shaped. Vessel trench is detected using the algorithm that is used for the measurement of curvature [12]. The Eigenvalues of the Hessian matrix be $\lambda'_{r,t}$ and $\lambda''_{r,t}$ at any pixel location (r,t) . Therefore, the principal curvature corresponding to minimum and maximum can be expressed as,

$$\hat{\lambda}_{r,t}^- = \min\{\lambda'_{r,t}, \lambda''_{r,t}\} \quad (5)$$

$$\hat{\lambda}_{r,t}^+ = \max\{\lambda'_{r,t}, \lambda''_{r,t}\} \quad (6)$$

Let $U \times V$ be the size of the greyscale image, then the minimum and

maximum principal curvature matrices can be represented as,

$$M^- = [\hat{\lambda}_{r,t}^-]_{U \times V} \quad (7)$$

$$M^+ = [\hat{\lambda}_{r,t}^+]_{U \times V} \quad (8)$$

The matrix for adaptive principal curvature can be expressed as, $M = \frac{1}{2}(M^+ + M^-)$. However, the maximum principal curvature provides better performance to extract the blood vessels, so we consider the maximum principal curvature expressed by Eq. (11).

Step (iii): After segmentation, the image passes through two sections, one, excitation and bottleneck block, and, other, Convolutional Neural Network. The output is obtained individually for both paths.

Step (iv): While training, the section of the squeeze and excitation and bottleneck blocks is trained only on normal images and abnormal images. This serves as a memory module for the process. The next step involves, feeding the whole training the CNN with the entire training data, i.e. both abnormal and normal images.

Step (v): For the test image to be classified, it first passes through the excitation and bottleneck blocks, giving *output 1*, while the test image passes through CNN, giving *output 2*.

Step (vi): If both the *output 1* and *output 2* have classified the image as normal, then the image is classified as a normal image. If both the *output 1* and *output 2* has classified the image abnormal, then it is declared and displayed as abnormal. Else, the image is further processed using the CNN algorithm to obtain the classified result. The overview of the proposed system for output generation of diabetes detection is shown in Fig. 4.

To train the CNN architecture effectively, the training data is split into two groups, the ‘Normal’, and the ‘Abnormal’. The CNN takes an input mage of 336×448 and passes it through several layers of convolution and max-pooling. The first convolution layer contains 10 number of 9×9 filters preceding a 2×2 max-pooling layer and the next convolution layer contains 10 number of 6×6 filters followed by a 3×3 max-pooling layer. Following this is an FC layer of output size 2, batch normalization layer, softmax layer, and finally the classification layer which also uses ReLU activation. At a learning rate of 0.00001, training of the network is done for 20 epochs. Since the segmented images show the volume of blood vessels present, the model is trained in a way that if

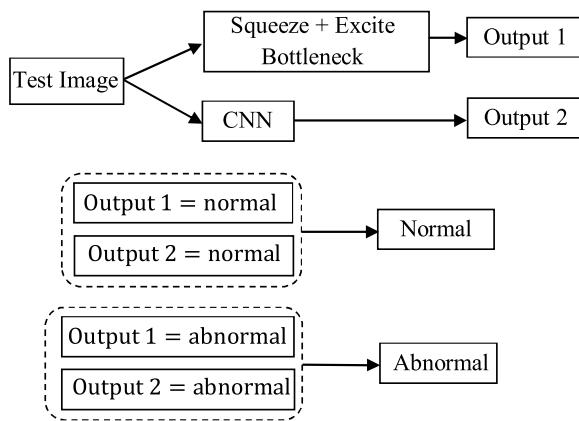


Fig. 4. Overview of output generated by the proposed method.

the segmented image shows an abnormal growth of blood vessels, then it is classified as abnormal. Whereas, if the segmented image shows no unusual growth of the blood vessels, then it is classified as normal. So, the feature that is used to decide the condition of the patient is the blood vessel structure present inside the retina. The central CNN is trained on both classes that consist of convolutional layers with an FC layer at the end to classify any input image to one of the two categories, ‘Normal’ or ‘Abnormal’. Only if the prediction of either memory module coincides with that of the main CNN, then the image of the fundus is classified into that particular category, else it is further processed till classification completes. The next section shows the experimental results and discussion of the proposed work.

4. Experimental outcomes and discussion

The proposed algorithm was assessed using the images taken from the dataset and the practical images obtained from the hospital.

4.1. Data collection and processing

Our data contains high-resolution fundus (HRF) images [5] that are distributed into two groups, i.e. healthy, and diabetic retinopathy, each comprising twenty images. For testing purposes, eye hospital Sankara Nethralaya has provided us with 80 patient images from their patient database. Also, we have used the standard Diabetic Retinopathy Dataset [33] (DIARETDB1), which contains the fundus images of 89 patients. The DIARETDB1 dataset contains the retina fundus that was acquired at the 50° field of view. The dataset contains 84 DR images and 5 normal images each having a size of 1500 × 1152 pixels. The dataset images are annotated by 4 medical experts for the presence of microaneurysms, soft exudates, hard exudates, and hemorrhages [34,35]. The experimental

results were evaluated based on k-fold cross validation where k is chosen as 5. The proposed architecture was trained using $k - 1 = 4$ folds and the model was tested on the remaining 5th fold. The performance metric was recorded for all the k tests (iterations) and the average score is considered as the performance of the proposed model. The performance was measured without data augmentation and with the data augmentation process. The data augmentation includes flipping such as horizontal and vertical flipping and rotation such as 45°, 90°, 135°, 180°, 225°, 270° and 315° rotation. We have done pre-processing such as cropping the background and resized the image to 512 × 512 and converted to grayscale to make the image to be uniform. Figs. 5 and 6 shows the sample images from the dataset DIARETDB1 and hospital respectively.

4.2. Experimental outcomes

The proposed system for the detection of Diabetes based on the proliferation of optical vasculature has been implemented using the tool MATLAB 2018b with Computer Vision System Toolbox. In our work, the maximum principal curvatures are determined using the ‘lambda’ function in MATLAB which performs region growing along the direction of their eigenvectors leading to the segmentation of the vasculature. When an image is fed in, it first goes through the steps of pre-processing, including resizing, converting to greyscale, and getting segmented output by the method of maximum principal curvature, which branches out the blood vessels. The standard deviation of the segmentation result varies between 0.43 to 3.79. To overcome any deficiencies in the data, utilizing the data available fully, augmentation techniques of data were carried out which included flipping and rotation. The image is then ready to be feed into both the memory block and the CNN. The central CNN architecture involves, two max-pooling layers, a couple of convolution layers, two FC layers, ReLU activation layers, with a softmax layer concluding in a classification layer. The memory block which includes the squeeze plus excitation and the bottleneck block, too involves convolution layers, batch Normalization layers, ReLU layers, and max-pooling layers. It takes twenty epochs, 0.0001 learning rate, and stochastic gradient descent with momentum optimizer.

While training, the whole training data is divided in the ratio of 4:1 as prediction and validation data. The validation data is further utilized to calculate the training accuracy and also other performance parameters. When a test image is taken and fed into the network, the image undergoes the segmentation process, and further, it was classified as an ‘Abnormal’ or ‘Normal’ image.

Fig. 7 shows some of the sample segmentation results for the dataset images DIARETDB1. The proposed method segments the fine blood vessels which can be best suited for classification. The segmentation process does not skip fine details present in the fundus image to the branching of blood vessels. Also, Fig. 8 shows the segmentation results

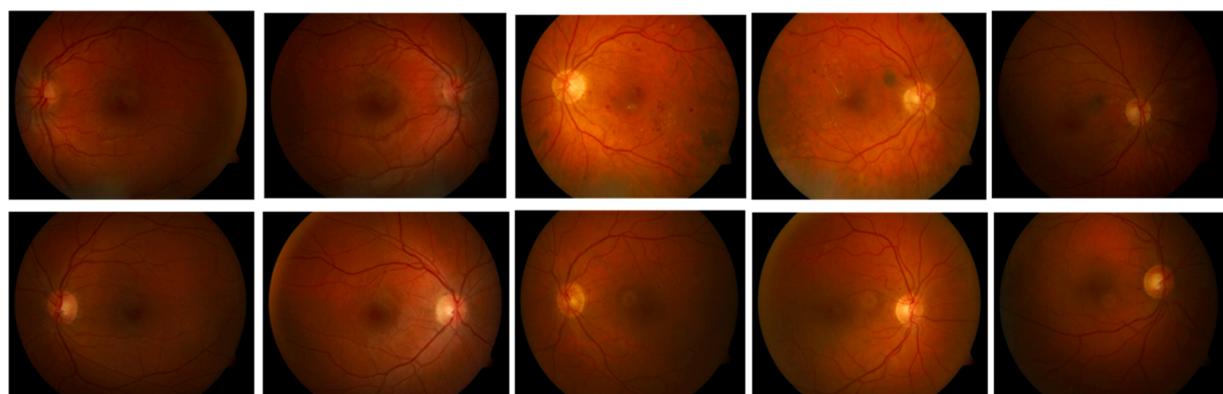


Fig. 5. Sample test images from Standard Diabetic Retinopathy Dataset (DIARETDB1) [30].

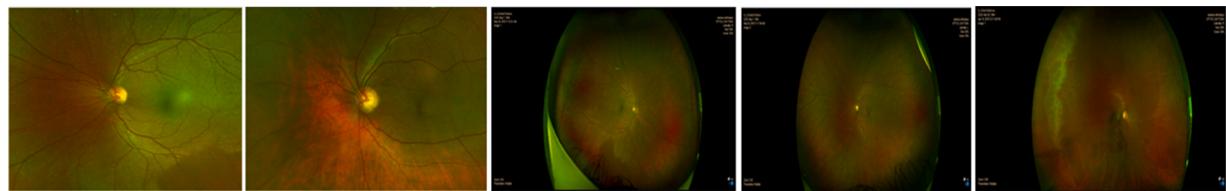


Fig. 6. Sample test images obtained from the Hospital.

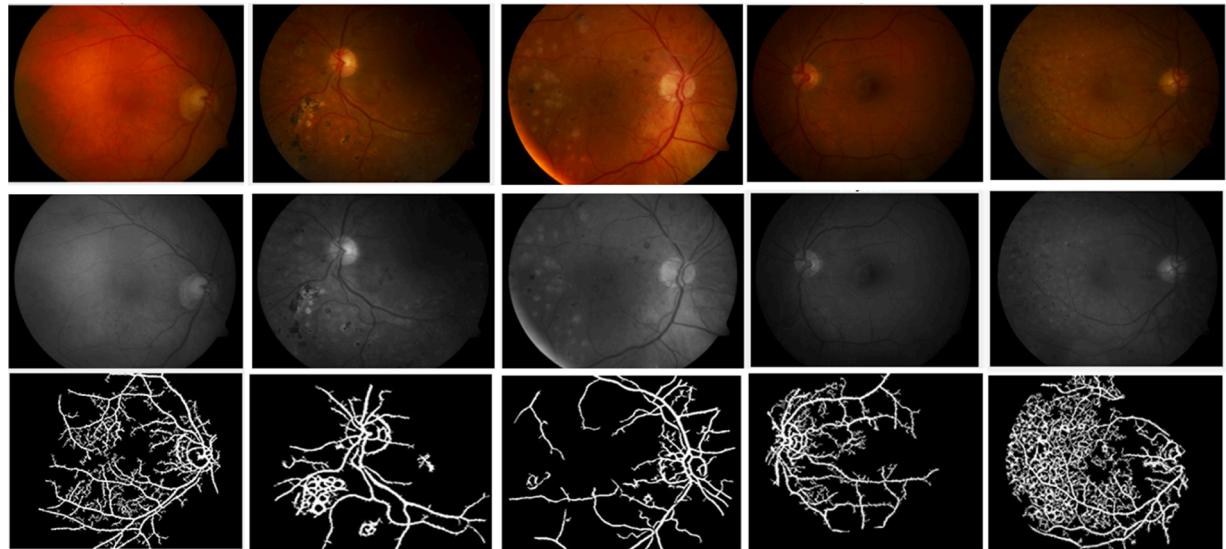


Fig. 7. Some of the segmentation results for the dataset DIARETDB1 images that are classified as abnormal (*First row*: Input image, *Second row*: Greyscale converted, *Third row*: Segmented output).

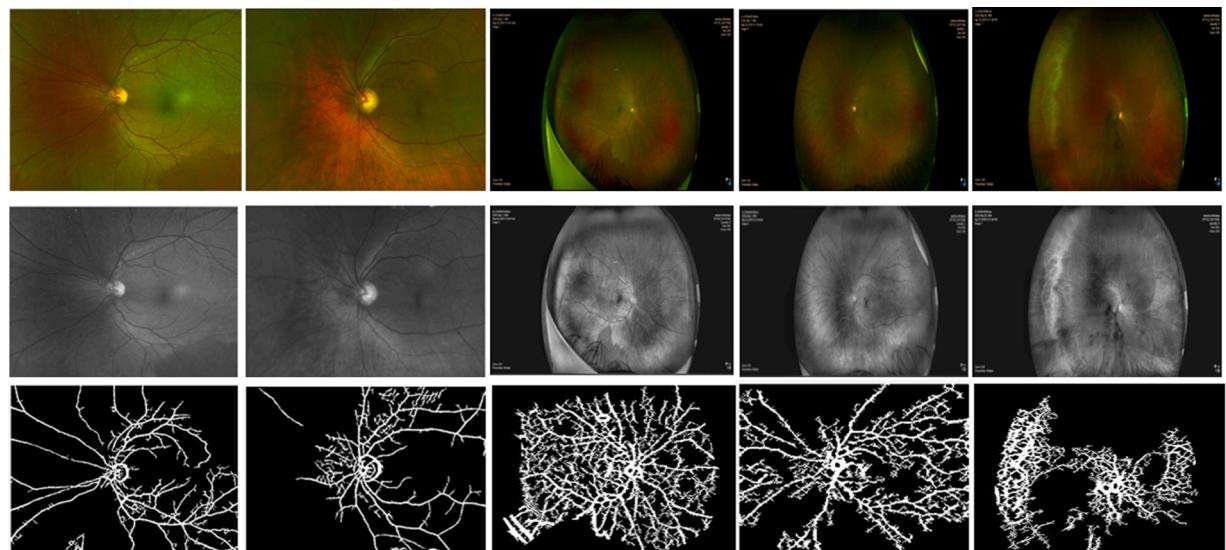


Fig. 8. Some of the segmentation results for the hospital images (*First row*: Input image, *Second row*: Greyscale converted, *Third row*: Segmented output).

for the hospital images. Even though the input image highly varies from the dataset images, the proposed segmentation process also obtains the fine details present in the hospital images without missing any major blood vessels.

The segmentation result of the proposed method highly depends on the greyscale conversion, the traditional conversion using Eq. (4) that misses the fine details as shown in Fig. 9 (c). But the greyscale conversion by providing high weightage to R channel using Eq. (5) provides

fine details at the output as shown in Fig. 9(e).

The performance of the algorithm was validated using the parameters such as precision, recall, sensitivity, and accuracy. The accuracy determines the overall correctness of the classifier, where the precision indicates the authenticity of the predicted positives. While the proportion of actual DR cases/ normal that are properly identified as such is called Sensitivity, that of actual non-DR/ abnormal that are properly identified as such is Specificity. The precision, specificity, recall, and

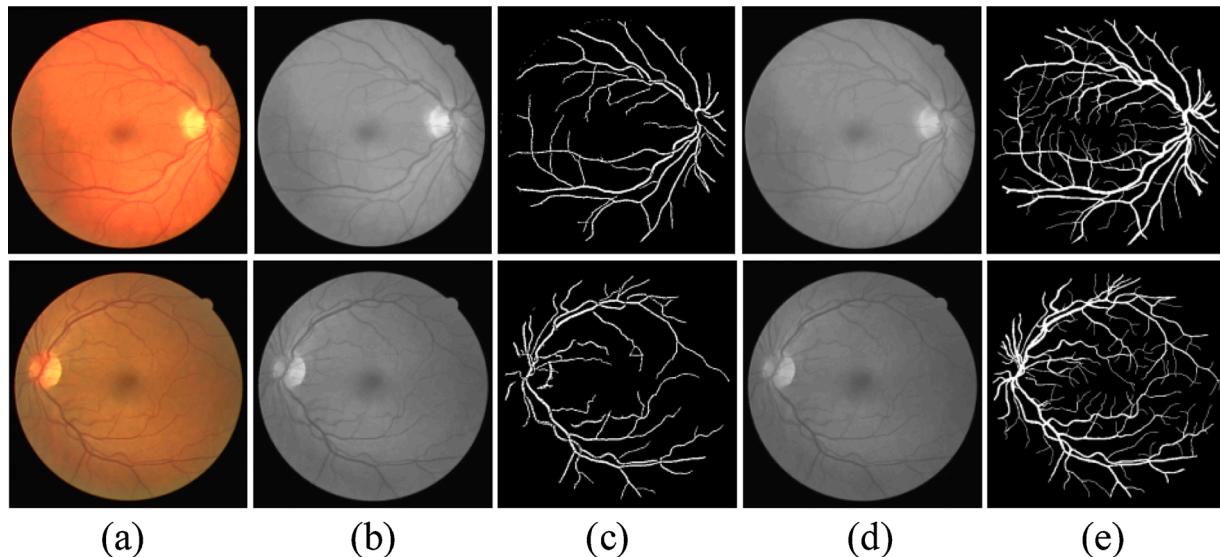


Fig. 9. Some segmented results with different Greyscale conversion (a) Input image, (b) Greyscale converted using Eq. (4),(c) Segmented output with for greyscale conversion using Eq. (4),(d) Greyscale converted using Eq. (5),(e) Segmented output with for greyscale conversion using Eq. (5).

accuracy can be estimated using the relation (12), (13), (14), and (15) respectively as,

$$\text{Precision} = \frac{T_{pos}}{T_{pos} + F_{pos}} \quad (9)$$

$$\text{Specificity} = \frac{T_{neg}}{T_{neg} + F_{pos}} \quad (10)$$

$$\text{Sensitivity} / \text{Recall} = \frac{T_{pos}}{T_{pos} + F_{neg}} \quad (11)$$

$$\text{Accuracy} = \frac{T_{pos} + T_{neg}}{T_{pos} + T_{neg} + F_{pos} + F_{neg}} \quad (12)$$

Here F_{pos} , F_{neg} T_{neg} , and T_{pos} represent false positive, false negative, true negative, and true positive, respectively. The performance of our algorithm with and without data augmentation was compared with the traditional algorithms such as Deep Learning [36], Red Lesions [37], Multi-Sieving [38], Ensemble Approach [39], and circular Hough-CNN [24] using the parameters like precision, specificity, recall and accuracy as shown in Table 1.

4.3. Experimental discussions

The precision, specificity, recall, and accuracy of the proposed method were estimated separately for both the dataset image and the hospital image. The precision, specificity, recall, and accuracy of the proposed algorithm, when tested using DIARETDB1 dataset images, provides a higher value when compared to the hospital images. For the DIARETDB1 dataset, the precision, specificity, recall, and accuracy of the proposed system without data augmentation were estimated to be 96.4 %, 97.2 %, 98.3 %, and 96.92 % respectively. For the same DIARETDB1 dataset the precision, specificity, recall, and accuracy increase to 97.2 %, 98.2 %, 99.6 %, and 98.7 % respectively with data augmentation. For the hospital dataset the precision, specificity, recall, and accuracy is estimated to be 96.1 %, 97.6 %, 97.2 %, and 97.4 % respectively with data augmentation. When comparing the performance of our method with the traditional method using the DIARETDB1 dataset, the proposed method with data augmentation has an improvement of 15.2 % in precision in juxtaposition to the scheme Red lesions [37]. The Specificity of the proposed method improves by 0.5 % when compared with the circular Hough-CNN [24]. The recall and accuracy of the proposed method increase only by 0.004 and 0.17 % respectively when compared to the scheme circular Hough-CNN [24]. Fig. 10 shows the graphical comparison of precision, specificity, recall, and accuracy of the proposed method with the traditional schemes.

To avoid the overfitting in the DIARETDB1 dataset, data augmentation is used that includes flipping such as horizontal and vertical flipping and rotation such as 45° , 90° , 135° , 180° , 225° , 270° and 315° rotation. The usage of this data augmentation method avoids overfitting.

Table 1
Comparison of proposed and traditional methods in terms of performance.

| Schemes | Precision (%) | Specificity (%) | Recall | Accuracy (%) |
|--|---------------|-----------------|--------|--------------|
| Deep Learning [36] | 79 | 82 | 0.84 | 86.1 |
| Red Lesions [37] | 82 | 83.1 | 0.881 | 92 |
| Multi-Sieving [38] | 81.5 | 92.1 | 0.978 | 96.1 |
| Ensemble Approach [39] | 80.3 | 85.5 | 0.881 | 86.9 |
| Circular Hough – CNN [24] | – | 97.7 | 0.992 | 98.53 |
| Proposed (DIARETDB1) without data augmentation | 96.4 | 97.2 | 0.983 | 96.92 |
| Proposed (Hospital) without data augmentation | 95.32 | 96.3 | 0.965 | 95.42 |
| Proposed (DIARETDB1) with data augmentation | 97.2 | 98.2 | 0.996 | 98.7 |
| Proposed (Hospital) with data augmentation | 96.1 | 97.6 | 0.972 | 97.4 |

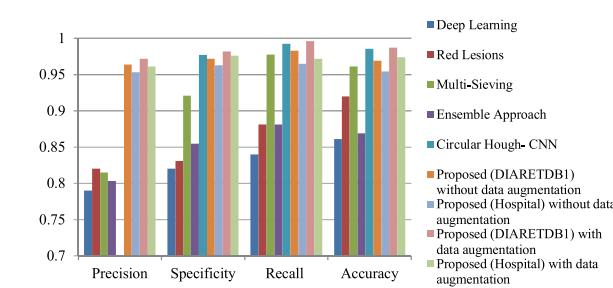


Fig. 10. Performance Comparison of proposed methods with the traditional classification schemes.

In the case of the hospital dataset, both the data augmentation and 20 % dropout of neurons at the hidden layer were used to avoid overfitting. Also, we have used 80 % data for the training set, and 20 % of data for validation to test the overfitting and found that no overfitting occurs with a learning rate of 0.00001 for 20 epochs. The area under the curve is observed from the ROC curve plot obtained from plotting the true positive rate against the false-positive rate. For the proposed methods for the dataset 'DIARETDB1' and hospital images, the ROC curve shows a better performance than the traditional schemes. The hospital images and the images from the DIARETDB1 dataset almost show similar ROC characteristics as shown in Fig. 11. The average classification for the proposed algorithm is estimated as 73 milliseconds.

Table 2 and Fig. 12 show the comparison of Precision, Specificity, and accuracy for each fold with its average value. The precision, specificity, and accuracy obtained on each fold are found to be closer to its average value. Thus, for each fold, there is not much variation in the performance metrics in the k-fold cross-validation.

Accuracy and its variance for DIARETDB1 and hospital dataset are depicted in Table 3 and Fig. 13. When learning 20 epochs, the DIARETDB1 has an accuracy of 98.7 % which is higher than the accuracy of the hospital dataset that has the accuracy of 97.4 %. Especially in the 1st epoch, the DIARETDB1 has an accuracy of 91.3 % and the hospital dataset has an accuracy of 90.2 %. The variance is high at the 11th epoch for the DIARETDB1 dataset where the accuracy is 96.12 %. For the Hospital dataset, the variance is maximum at the 14th epoch, which has an accuracy of 95.23 %. The next section shows the conclusion of the proposed work.

5. Conclusion

This paper proposed a Diabetic retinopathy detection and classification algorithm that uses CNN. Blood clots/blobs, exudates, and abnormal growth of blood vessels can be observed in the fundus of patients suffering from DR. This method initially pre-processes the images and the branching blood vessels are extracted through the segmentation process. Maximum principal curvature, which utilizes the maximum Eigenvalues of the Hessian matrix, has been applied to extract the branching blood vessels post. The adaptive histogram equalization and the morphological opening is performed to enhance and eliminate falsely segmented regions. Further, two sub-networks of neural networks are created excitation and bottleneck, and CNN. If both the sub-networks classify the same label as output, then the image is classified as that label. The experimental results are evaluated using the DIARETDB1 dataset and the images obtained from the hospital. The precision, specificity, recall, and accuracy of the proposed method with data

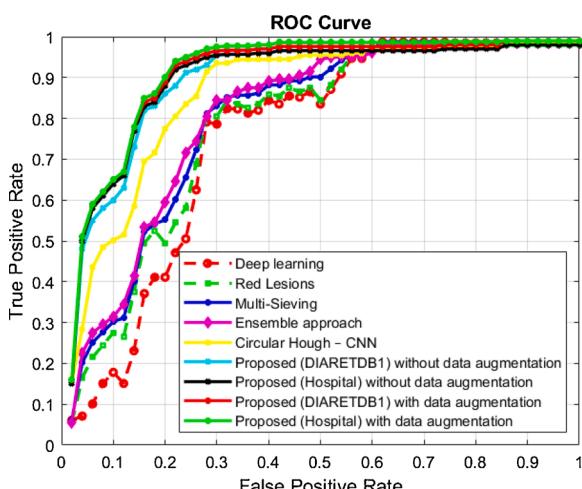


Fig. 11. Comparison of ROC for the proposed and traditional methods.

Table 2

Comparison of Precision, Specificity, and Accuracy on each fold.

| Testing fold | Precision | Specificity | Accuracy |
|--------------|-----------|-------------|----------|
| 1st fold | 95.74 | 98.12 | 98.54 |
| 2nd fold | 95.12 | 96.34 | 97.98 |
| 3rd fold | 96.33 | 97.19 | 95.86 |
| 4th fold | 95.98 | 97.56 | 96.21 |
| 5th fold | 97.42 | 98.77 | 98.52 |
| Average | 96.1 | 97.6 | 97.4 |

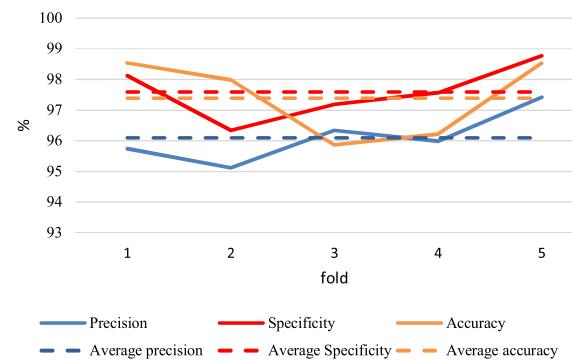


Fig. 12. Comparison of Precision, Specificity, and Accuracy on each fold.

Table 3

Accuracy and dispersion of DIARETDB1 and Hospital dataset.

| Epoch | DIARETDB1 | | Hospital | |
|-------|-----------|------------|----------|------------|
| | Accuracy | Dispersion | Accuracy | Dispersion |
| 1 | 91.3 | 0.12 | 90.2 | 0.37 |
| 2 | 91.9 | 0.25 | 90.52 | 0.47 |
| 3 | 92.42 | 0.52 | 90.83 | 0.62 |
| 4 | 92.78 | 0.93 | 91.14 | 0.78 |
| 5 | 93.17 | 0.54 | 91.36 | 0.63 |
| 6 | 93.62 | 0.92 | 91.57 | 0.47 |
| 7 | 93.91 | 0.14 | 91.87 | 0.76 |
| 8 | 94.12 | 0.31 | 92.34 | 0.53 |
| 9 | 94.43 | 0.76 | 92.67 | 0.86 |
| 10 | 94.71 | 2.31 | 92.91 | 0.96 |
| 11 | 96.12 | 2.62 | 93.51 | 1.43 |
| 12 | 97.32 | 0.82 | 93.81 | 1.05 |
| 13 | 97.54 | 0.92 | 94.12 | 2.04 |
| 14 | 97.83 | 0.41 | 95.23 | 2.53 |
| 15 | 97.97 | 0.42 | 96.02 | 0.62 |
| 16 | 98.21 | 0.41 | 96.54 | 0.81 |
| 17 | 98.45 | 0.63 | 96.89 | 0.23 |
| 18 | 98.61 | 0.75 | 97.2 | 0.25 |
| 19 | 98.7 | 0.43 | 97.4 | 0.93 |
| 20 | 98.7 | 0.43 | 97.4 | 0.93 |

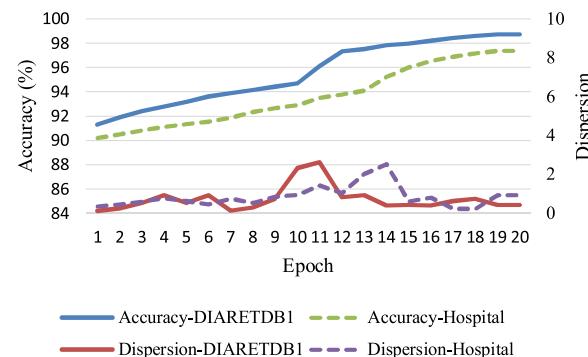


Fig. 13. Comparison of Accuracy and dispersion for DIARETDB1 and Hospital dataset.

augmentation for the DIARETDB1 dataset images was estimated to be 97.2 %, 98.2 %, 99.6 %, and 98.7 % respectively, which is higher than the traditional schemes. In the future, the classification can be further diversified based on the period of suffering. If the image is categorized as ‘Abnormal’ i.e. the patient has been detected with DR in the fundus image, additional processing can be done to classify it, based on how long the patient has been facing the problem. This could help to administer the treatment given to each patient and regulate the dosage of medication.

CRediT authorship contribution statement

Sraddha Das, Krity Kharbanda: Software implementation, Visualization

Suchetha M: Supervision, Conceptualization, Methodology, Data curation, Analysing, Investigation

Rajiv Raman: Investigation, Testing and validation

Edwin Dhas D: Visualization, Writing- Reviewing and Editing

Declaration of Competing Interest

The authors declare no conflict of interest.

References

- [1] N.C. Santosh, Y. Radhika, Optimized maximum principal curvatures based segmentation of blood vessels from retinal images, *Biomed. Res.* (2019), <https://doi.org/10.35841/biomedicalresearch.30-19-068>.
- [2] Gehad Hassan, Nashwa El-Bendary, Aboul Ella Hassani, Aly Fahmy, M. Shoeb, Vaclav Snasel, Retinal blood vessel segmentation approach based on mathematical morphology, *Procedia Comput. Sci.* 65 (2015) 612–622, <https://doi.org/10.1016/j.procs.2015.09.005>.
- [3] J.I. Orlando, E. Prokofyeva, M.B. Blaschko, A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images, *IEEE Trans. Biomed. Eng.* 64 (1) (2017) 16–27, <https://doi.org/10.1109/TBME.2016.2535311>.
- [4] R. Manjula Sri, J. Jyothirmai, D. Swetha, Analysis of retinal blood vessel segmentation in different types of diabetic retinopathy, *Int. J. Eng. Adv. Technol. (IJEAT)* (2019).
- [5] N. Memari, A.R. Ramli, M.I.B. Saripan, et al., Retinal blood vessel segmentation by using matched filtering and fuzzy C-means clustering with integrated level set method for diabetic retinopathy assessment, *J. Med. Biol. Eng.* (2019) 713–731, <https://doi.org/10.1007/s40846-018-0454-2>.
- [6] Attila Budai, Rüdiger Bock, Andreas Maier, Joachim Hornegger, Georg Michelson, Robust vessel segmentation in fundus images, *Int. J. Biomed. Imaging* (2013), 154860, <https://doi.org/10.1155/2013/154860>.
- [7] Z. Tian, L. Liu, Z. Zhang, B. Fei, Superpixel-based segmentation for 3D prostate MR images, *IEEE Trans. Med. Imaging* 35 (3) (2015) 791–801.
- [8] D.C.T. Nguyen, S. Benamou, M. Mignotte, F. Lavoie, Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images, *Med. Image Anal.* 48 (2018) 58–74.
- [9] Y.L. Huang, D.R. Chen, Watershed segmentation for breast tumor in 2-D sonography, *Ultrasound Med. Biol.* 30 (5) (2004) 625–632.
- [10] H. Masoumi, A. Behrad, M.A. Pourmina, A. Roosta, Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network, *Biomed. Signal Process. Control* 7 (5) (2012) 429–437.
- [11] M. Ciechlewski, An edge-based active contour model using an inflation/deflation force with a damping coefficient, *Expert Syst. Appl.* 44 (2016) 22–36.
- [12] M. Ciechlewski, J.H. Spodnik, Semi-automatic corpus callosum segmentation and 3d visualization using active contour methods, *Symmetry* 10 (11) (2018) 589.
- [13] D. Lokuarachchi, K. Gunarathna, L. Muthumal, T. Gamage, Automated detection of exudates in retinal images, in: *IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA)*, Penang, Malaysia, 2019, pp. 43–47, <https://doi.org/10.1109/CSPA.2019.8696052>.
- [14] R.J. Chalakkal, W.H. Abdulla, S.S. Thulaseedharan, Automatic detection and segmentation of optic disc and fovea in retinal images, *IET Image Process.* 12 (11) (2018) 2100–2110, <https://doi.org/10.1049/iet-ipr.2018.5666>, 11.
- [15] S. Selvaperumal, Ramasubramanian Bhoopalan, An efficient approach for the automatic detection of hemorrhages in colour retinal images, *IET Image Process.* (2018) 12, <https://doi.org/10.1049/iet-ipr.2017.1036>.
- [16] A. Benzamin, C. Chakraborty, Detection of Hard exudates in retinal fundus images using deep learning, in: 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 2018, pp. 465–469, <https://doi.org/10.1109/ICIEV.2018.8641016>.
- [17] S. Yu, D. Xiao, Y. Kanagasingam, Exudate detection for diabetic retinopathy with convolutional neural networks, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, 2017, pp. 1744–1747, <https://doi.org/10.1109/EMBC.2017.8037180>.
- [18] Shailesh Kumar, Abhinav Adarsh, Basant Kumar, Amit Singh, An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation, *Opt. Laser Technol.* (2020), <https://doi.org/10.1016/j.optlastec.2019.105815>.
- [19] N. Yalçın, S. Alver, N. Uluhatun, Classification of retinal images with deep learning for early detection of diabetic retinopathy disease, in: 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1–4, <https://doi.org/10.1109/SIU.2018.8404369>.
- [20] Payal Shah, Divyansh K. Mishra, Mahesh P. Shanmugam, Bindiya Doshi, Hariprasad Jayaraj, Rajesh Ramanjulu, Validation of Deep Convolutional Neural Network-based algorithm for detection of diabetic retinopathy – artificial intelligence versus clinician for screening, *Indian J. Ophthalmol.* 68 (2020) 398, https://doi.org/10.4103/ijo.IJO_966_19.
- [21] D.Y. Carson Lam, M. Guo, T. Lindsey, Automated detection of diabetic retinopathy using deep learning, *AMIA Summits Transl. Sci. Proc.* 2018 (2018) 147 [PMC free article].
- [22] M. Andonová, J. Pavlovičová, S. Kajan, M. Oravec, V. Kurilová, Diabetic Retinopathy Screening Based on CNN, *International Symposium ELMAR*, Zadar, 2017, pp. 51–54, <https://doi.org/10.23919/ELMAR.2017.8124433>.
- [23] P. Prentasić, S. Lončarić, Detection of exudates in fundus photographs using convolutional neural networks, in: 9th International Symposium on Image and Signal Processing and Analysis (ISPA), Zagreb, 2015, pp. 188–192, <https://doi.org/10.1109/ISPA.2015.7306056>.
- [24] Kemal Adem, Exudate detection for diabetic retinopathy with circular hough transformation and convolutional neural networks, *Expert Syst. Appl.* (2018) 114, <https://doi.org/10.1016/j.eswa.2018.07.053>.
- [25] K. Lyu, Y. Li, Z. Zhang, Attention-aware multi-task convolutional neural networks, *IEEE Trans. Image Process.* 29 (2020) 1867–1878, <https://doi.org/10.1109/TIP.2019.2944522>.
- [26] Thomas A. Siji, Titus Geeverghese, Design of a portable retinal imaging module with automatic abnormality detection, *Biomed. Signal Process. Control* 60 (2020), <https://doi.org/10.1016/j.bspc.2020.101962>.
- [27] D.J. Hemanth, O. Deperioglu, U. Kose, An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network, *Neural Comput. Appl.* 32 (2020) 707–721, <https://doi.org/10.1007/s00521-018-03974-0>.
- [28] S. Yang, C. Yang, D. Feng, X. Hao, M. Wang, One-dimensional deep attention convolution network (ODACN) for signals classification, *IEEE Access* 8 (2020) 2804–2812, <https://doi.org/10.1109/ACCESS.2019.2958131>.
- [29] W. Hu, H. Li, L. Pan, W. Li, R. Tao, Q. Du, Spatial-Spectral feature extraction via deep ConvLSTM neural networks for hyperspectral image classification, *IEEE Trans. Geosci. Remote. Sens.* 58 (June 6) (2020) 4237–4250, <https://doi.org/10.1109/TGRS.2019.2961947>.
- [30] Qi Zhao, Jiahui Liu, Boxue Zhang, Shuchang Lyu, Nauman Raoof, Wenquan Feng, Interpretable Relative squeezing bottleneck design for compact convolutional neural networks model, *Image Vis. Comput.* 89 (2019), <https://doi.org/10.1016/j.imavis.2019.06.006>.
- [31] Jun Gu, Xuewen Sun, Yue Zhang, Kun Fu, Lei Wang, Deep residual squeeze and excitation network for remote sensing image super-resolution, *Remote Sens.* 11 (1817) 2019, <https://doi.org/10.3390/rs11151817>.
- [32] S. Lahmiri, Hybrid deep learning convolutional neural networks and optimal nonlinear support vector machine to detect presence of hemorrhage in retina, *Biomed. Signal Process. Control* 60 (2020), <https://doi.org/10.1016/j.bspc.2020.101978>.
- [33] L. Qiao, Y. Zhu, H. Zhou, Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms, *IEEE Access* (2020), <https://doi.org/10.1109/ACCESS.2020.2993937>.
- [34] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, H. Kang, Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening, *Inform. Sci.* 501 (2019) 511–522.
- [35] W.L. Alyoubi, W.M. Shalash, M.F. Abulkhair, Diabetic retinopathy detection through deep learning techniques: a review, *Inform. Med. Unlocked* (2020), 100377.
- [36] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, J. Pietilä, DIARETDB1 diabetic retinopathy database and evaluation protocol, *Technical report (PDF)* (2021).
- [37] K.M. Adal, P.G. Van Etten, J.P. Martinez, K.W. Rouwen, K.A. Vermeer, L.J. van Vliet, An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images, *IEEE Trans. Biomed. Eng.* 65 (6) (2017) 1382–1390, <https://doi.org/10.1109/TBME.2017.2752701>.
- [38] L. Dai, R. Fang, H. Li, X. Hou, B. Sheng, Q. Wu, W. Jia, Clinical report guided retinal microaneurysm detection with multi-sieving deep learning, *IEEE Trans. Med. Imaging* 37 (5) (2018) 1149–1161, <https://doi.org/10.1109/TMI.2018.2794988>.
- [39] S. Qummar, F.G. Khan, S. Shah, A. Khan, S. Shamshirband, Z.U. Rehman, I.A. Khan, W. Jadoon, A deep learning ensemble approach for diabetic retinopathy detection, *IEEE Access* 7 (2019) 150530–150539, <https://doi.org/10.1109/ACCESS.2019.2947484>.