A New Early Stage Diabetic Retinopathy Diagnosis Model Using Deep Convolutional Neural Networks and Principal Component Analysis

Mali Mohammedhasan*, Harun Uğuz

Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Selçuk Üniversitesi, Konya 42130, Turkey

Corresponding Author Email: mali.mohammedhasan@selcuk.edu.tr

ABSTRACT

Diabetic retinopathy (DR) is a disease of the retina, which leads over time to vision problems such retinal detachment, vitreous hemorrhage, glaucoma, and in worse cases leads to blindness, which can initially be controlled by periodic DR-screening. Early diagnosis will lead to greater control of the disease, whereas performing retinal examinations on all diabetic patients is an unattainable need, as diabetes is a chronic disease and its global prevalence has been steadily increasing over the past few decades. According to recent World Health Organization statistics, about 422 million people worldwide have diabetes, the majority living in low-and middle-income countries. This paper proposes a new strategy that brings the strength of convolutional neural networks (CNNs) to the diagnosis of DR. Coupled with using principal component analysis (PCA) that performs dimension reduction to improve the diagnostic accuracy, the proposed model exploiting edge-preserving guided image filtering (E-GIF) that performs as a contrast enhancement mechanism, and in addition to smoothing low gradient areas, it also accentuates strong edges. Diabetic retinopathy causes progressive damage to the blood vessels in the retina to the extent that it leaves traces and lesions in the tissues of the retina. These lesions appear in the form of edges and when processing retinal images, we seek to accentuate these edges to enable better diagnosis of diabetic retinopathy symptoms. A new CNN architecture with residual connections is used, which performs very well in diagnosing DR. The proposed model is named with RUnet-PCA: Residual U-net Deep CNN with Principal Component Analysis. The well-known AlexNet, VggNet-s, VggNet-16, VggNet-19, GoogleNet, and ResNet models were adopted for comparison with the proposed model. Publicly available Kaggle dataset was employed for training exploring the DR diagnosis accuracy. Experimental results show that the proposed RUnet-PCA model achieved a diagnosis accuracy of 98.44% and it was extremely robust and promising in comparison to other diagnosis methods.

1. INTRODUCTION

Diabetic retinopathy (DR) is a state of disturbance that arises in the diabetes mellitus-induced retina [1]. It is estimated that there are around 93 million people living with DR, and around 28 million of DR-patients worldwide have suffered from vision loss for this cause [2]. Even though several considerations are influential in DR, metabolic end-products triggered by elevated blood sugar levels seem to be the most important aspect. The weakening of the walls of arteries causes bubbles, blockages and leakages throughout the veins [3]. It is very crucial to monitor retinal vascular blood flow for controlling disease symptoms before it spreads significantly. Accordingly, the information collected by retinal analysis indicate the activity of diabetes. Based on the findings of an extensive analysis carried out in the United States, the existence of DR brings a risk of death between 34% and 89% [4, 5].

Symptoms of diabetic retinopathy at the fundus of the eye include micro-aneurysm (MA), micro-hemorrhage (MH), hard exudate (HE), soft exudate (SE) and neo-vascularisation [6]. A several number of research studies have concentrated on the detection of these symptoms by the automated detecting systems [7-9]. Exudates are the initial phase of diabetic retinopathy, due to the high amount of lipoproteins that have leaked from the arteries and veins in people with diabetes. Early diagnosis is essential as these symptoms is mostly associated with visual loss [10]. In contrast to micro-hemorrhages, it demands more time to regain health by therapy. More and more articles have lately been presented in the literature on the diagnosis of such symptoms at the initial phases of DR [8, 10-12]. A normal and abnormal retinal sample is shown in Figure 1.

Figure 1. Sample retinal images from Kaggle dataset for the subject (337). The (left) shows clean retinal image, while the (right) shows DR-infected retinal image

As illustrated in Figure 1, the distinctions between normal retina and abnormal retina, with yellow exudate and hemorrhage, can be noted. The main objective of this study is to improve the efficiency of DR automatic diagnosis systems. It proposes a new early stage diabetic retinopathy diagnosis mechanism which exploits the role of deep convolutional neural networks in classifying retinal images.

The main contribution of the paper that it brings a new CNN architecture for DR detection, which is influenced by the well-known U-network with the inclusion of using residual connections to transfer information between layers, which showed superior performance over traditional methods based on feature extraction. Furthermore, principal component analysis (PCA) is used to performs dimension reduction to improve the diagnostic accuracy [13-15]. As well guided image filtering (GIF) is used to perform as a contrast enhancement for retinal images that not only smooths low gradients, furthermore preserves solid edges. GIF accentuates the edges (traces and lesions in the tissues) of the retina to enable better diagnosis of diabetic retinopathy symptoms. Finally, data augmentation technique is provided, which also improves the performance of the proposed model.

The paper layout is organized according to the following. The related work is presented in Section 2. Section 3 outlines materials and method. Section 4 explains the performance index and the output of the propose model. Finally, Section 5 summarizes the conclusions.

2. RELATED WORK

The research focus has been shifted to the development of effective CNN architectures after the popular CNN AlexNet propelled deep learning techniques into major functional applications [16]. Depth is the main consideration with respect to network efficiency. ResNet [17] incorporates identity skip-connections that eventuate deeper convolutional neural networks. DenseNet [18] and DPN [19] use layer-to-layer communication process modifications to improve deep neural network learning and representation functionalities. Veit believes that ResNet-like networks behave as fairly shallow network groups and enhancements in their efficiency are not because of deeper network architecture, but because of combinations of model [20]. Based on this point of view, the highway and inception [21, 22] architectures raise the number of sub-networks by expanding the depth of the network.

AlexNet is a spatial exploitation-based CNN architecture. Spatial exploitation means that the convolutional operation considers the neighborhood (correlation) of input pixels, therefore different levels of correlation can be explored by using different filter sizes. The major strengths associated with AlexNet are the extraction of low, mid and high-level features using large and small size filters on initial (5x5 and 11x11) and last layers (3x3), giving an idea of deep and wide CNN architectures, the introduction of regularization in CNNs, and starting parallel use of GPUs as an accelerator to deal with complex architectures. On the other hand, the major gaps associated with AlexNet are the inactive neurons in the first and second layers, and aliasing artifacts in the learned feature-maps due to large filter size [16].

ResNet, DenseNet, and Highway Networks are a multi-path-based CNN architecture. Multi-path means that shortcut paths provides the option to skip some layers. Different types of the shortcut connections used in literature are zero padded,

projection, dropout, 1x1 connections, etc.

The major strengths associated with ResNet are the use of identity-based skip connections to enable cross layer connectivity, the information flow gates are data independent and parameter free, and it can easily pass the signal in both directions, forward and backward. On the other hand, the major gaps associated with ResNet are that many layers may contribute very little or no information, also relearning of redundant feature-maps may happen [17].

The major strengths associated with DenseNet are the introducing of depth or cross-layer dimension, ensuring maximum data flow between the layers in the network, avoiding relearning of redundant feature-maps, and that low and high level both features are accessible to decision layers. On the other hand, the major gaps associated with DenseNet is the large increase in parameters due to increase in number of feature-maps at each layer [18].

The major strengths associated with Highway Networks is mitigating the limitations of deep networks by introducing cross layer connectivity. While the major gaps associated with Highway Networks is that gates are data dependent and thus may become parameter expensive [21, 22].

Earlier, it was assumed that to improve accuracy, the number of layers have to be increased. However, by increasing the number of layers, the vanishing gradient problem arises and training might get slow. So, the concept of widening a layer was also investigated. Inception is a width-based CNN architecture and its major strengths are that varying size filters inside inception module increases the output of the intermediate layers, and varying size filters are helpful to capture the diversity in high-detail images. On the other hand, the major gaps associated with Inception is the increase in space and time complexity.

The deep fusion network [23] also demonstrates that the deepest network in all the sections of the ensemble does not play a central role in enhancing the essential performance, but instead adds more layers to ensure the ensemble scale. Zhao introduces a new form of deep-fused network integrating two networks with a merging/running fusion technique [23]. It is more accurate than ResNet but has lower number of layers and equal number of parameters. ResNeXt [24] incorporates the ResNet residual block strategies completely in the inception network, as well as the split-transform-merge technique. Such homogenous multi-branch network has lower number of hyperparameters and smaller capacities as it only incorporates a clustered convolution technique, called "cardinality," in order to raise the dimension. The experiment findings indicate that, rising the cardinality is a more efficient method to enhancing network efficiency than deepening or widening it. Cross-channel correlations are drawn to form a composition of attributes, regardless of spatial structure [25] or together with regular convolutionary filters [26] that are shortcuts to 1×1 convolution operations.

Though most researches seek to minimize computational costs and simplify training challenges. Such researches are focused on the premise of a channel correlations that can be mapped as a combination of instance-agnostic features and local receptive fields for relationship among channels. On the other hand, Hu et al. [27] concentrates on the complex and non-linear interactions of channels that use global data sources. Such architecture design can be coupled with other well-known networks to promote and improve the learning and representational capabilities. PolyNet [28] discusses structural variety, an alternate aspect of network design that uses

different types of polynomial combinations for the generalization of residual inceptions.

3. MATERIALS AND METHOD

The proposed early stage diabetic retinopathy diagnosis model is named with RUnet-PCA as an abbreviation of Residual U-net Deep CNN with Principal Component Analysis. The whole flow of the proposed model involves four steps: image pre-processing, image enhancement using edge preserving GIF, data augmentation, and then diagnosis based on RUnet-PCA proposed model. Figure 2 shows the process diagram of proposed model. Figure 3 shows the process diagram of proposed DRUnet training model.
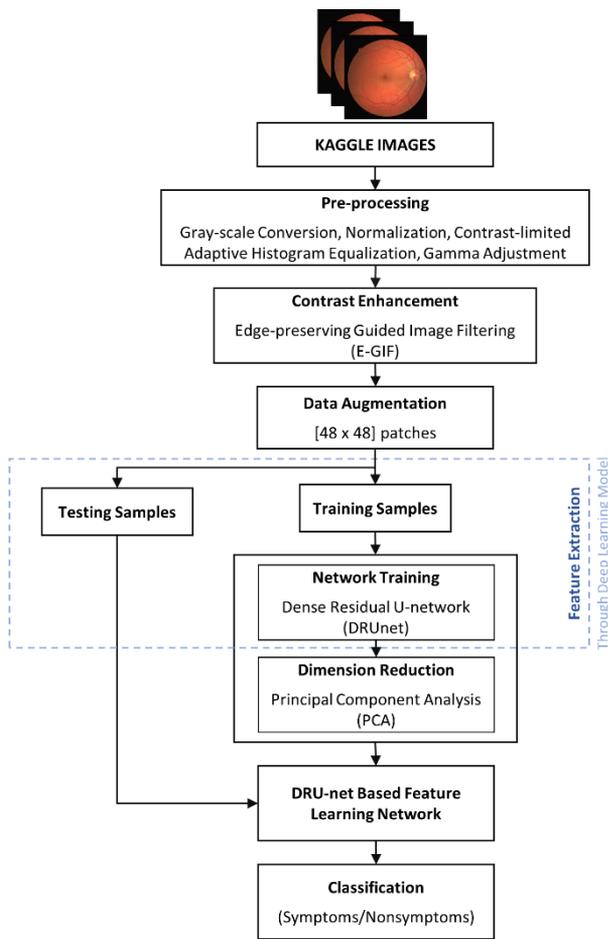
Figure 2. The process diagram of proposed model

Pseudo Code of The Algorithm:

STEP 1: use 2063 Kaggle images in its original format.
STEP 2: convert images to grayscale format, apply the transformations (standardization, CLAHE, and gamma adjustment).
STEP 3: apply edge-preserving guided image filtering to do contrast enhancement mechanism.
STEP 4: resized images [256 x 256], create 2063 x 256 x 256 data images.
STEP 5: increase data images by generating [48 x 48] dimensional patches by randomly choosing their centers within the complete original retinal image (data augmentation).

STEP 6: prepare samples, 60% for training and 40% for testing.
STEP 7: construct net of DRUnet.
STEP 8: train DRUnet network using training dataset to do feature extraction.
STEP 9: apply PCA to do dimension reduction of features.
STEP 10: evaluate and test the trained network using testing dataset, classify images using softmax function.
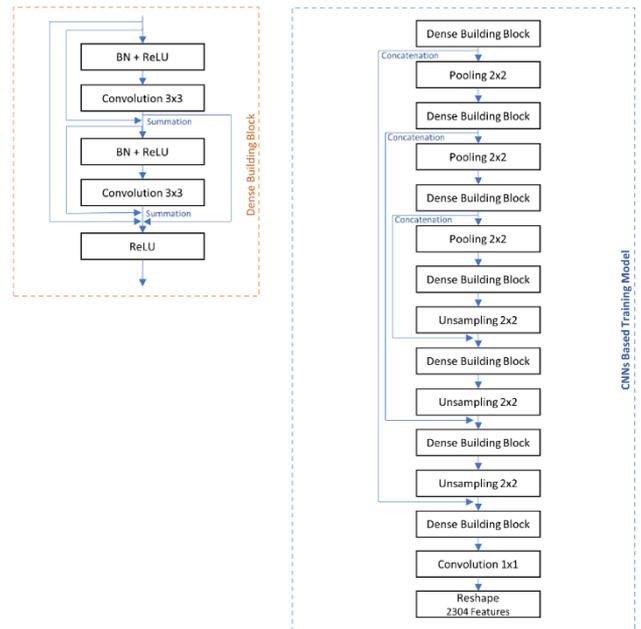
Figure 3. The process diagram of proposed DRUnet training model

3.1 Image pre-processing

Rahim et al. [29] explores the ability to improve classification efficiency through image pre-processing methods. In the proposed model, retinal images of Kaggle diabetic retinopathy dataset is firstly processed with a series of image pre-processing methods for promoting model classification while retaining as many of the original image features as possible.

In general, let us represent the series of used transformations mathematically as follows:

$$\tilde{x}_g = \text{grayscale}(x)$$
$$\tilde{x}_n = \frac{\tilde{x}_g - average(\tilde{x}_g)}{\max(\tilde{x}_g) - \min(\tilde{x}_g)} \text{ for mean normalization}$$
$$\tilde{x}_C = \text{CLAHE}(\tilde{x}_n)$$
$$\tilde{x}_{ga} = A\tilde{x}_C{}^{\gamma} \tag{1}$$

where, the input value $\tilde{x}_C$ is raised to the power γ and multiplied by the constant A. In the common case of A = 1, inputs and outputs are typically in the range 0–1.

Retinal images are processed with several pre-processing techniques prior to beginning network training, these techniques include the following transformation methods: gray-scale conversion transformation, standardization transformation, contrast-limited adaptive histogram equalization (CLAHE) transformation, and gamma adjustment transformation. Figure 4 (a-d) includes outputs of the pre-processing stage.

(a) The source image and the gray-scaled version
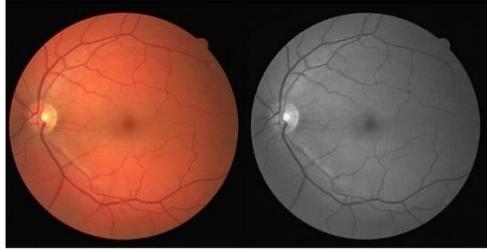
(b) The source image and normalized version
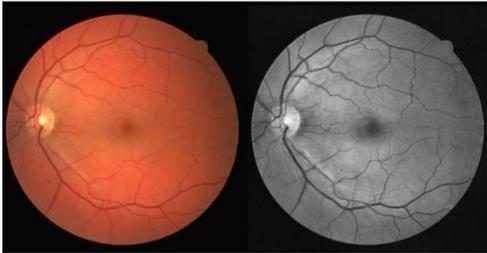
(c) The source image and CLAHE version
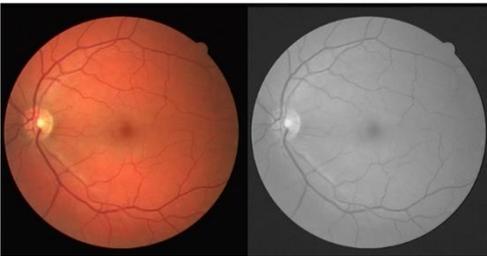
(d) The source and gama-corrected version [gama: 0.5]

Figure 4. Used pre-processing methods

3.2 Guided image filtering (edge-preserving filtering)

Edge-preserving guided image filtering (E-GIF) performs as a contrast enhancement mechanism, and in addition to smoothing low gradient areas, it also accentuates strong edges. Diabetic retinopathy causes progressive harm to the capillaries and arteries in the retina to the extent that it leaves traces and lesions in the tissues of the retina. These lesions appear in the form of edges and when processing retinal images, we seek to accentuate these edges to enable better diagnosis of DR symptoms. This is the motivation point for exploiting the characteristics of guided image filtering.

The guided filter practices an edge-preserving smoothing function on specific image, based on another image, called the guidance image, to perform the filtering [30]. The guiding image might be the image itself, or a modified version of it, or a totally different one [31].

If considering I_p and G_p are the intensity values at pixel p of the input and guided image, ω_k is the kernel window centered at pixel k. GIF is then expressed by:

$$GIF(I)_p = \frac{1}{\sum_{q \in \omega_k} W_{GIF_{pq}}(G)} \sum_{q \in \omega_k} W_{GIF_{pq}}(G) I_q \quad (2)$$

and the kernel weights function $W_{GIF_{pq}}(G)$ can be defined as:

$$W_{GIF_{pq}}(G) = \frac{1}{|\omega|^2} \sum_{k:(p,q) \in \omega_k} \left(1 + \frac{(G_p - \mu_k)(G_q - \mu_k)}{\sigma_k^2 + \varepsilon}\right) \quad (3)$$

where, μ_k and σ_k^2 are the mean and variance of guided image G in local window ω_k, $|\omega|$ is the number of pixels in this window. The term $(1 + \frac{(G_p - \mu_k)(G_q - \mu_k)}{\sigma_k^2 + \varepsilon})$ is the key to realizing the edge-preserving ability of GIF. When G_p and G_q are simultaneously on the same side of an edge (both are smaller or larger than the mean), the weight assigned to pixel q will be large. contrariwise, when they are on different sides (one is smaller and one is larger than be mean), a small weight will be assigned to pixel q. Computations [30] emphasize the normalization term in Eq. (2) equals 1. The filter kernel of GIF can be abbreviated as follows:

$$GIF(I)_p = \sum_{q \in \omega_k} W_{GIF_{pq}}(G) I_q \quad (4)$$

The level of smoothing/sharpening of GIF is controlled through the ε parameter.

The larger the value of ε parameter is, the more smoothly the filtered image would be. Figure 5 demonstrates the effect of guided image filtering on a sample retinal image by adjusting different values of ε parameter.
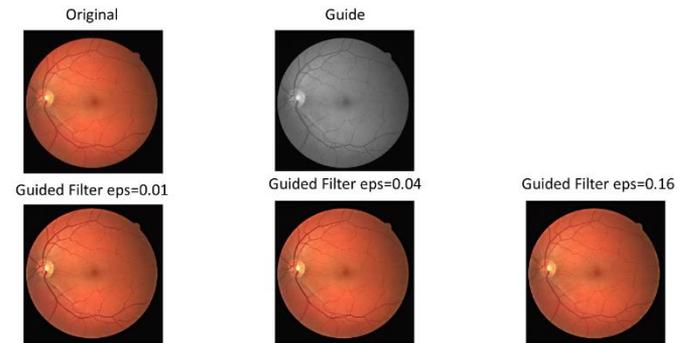
Figure 5. The effect of applying GIF on a sample retinal image

3.3 Principal component analysis

Principal Component Analysis (PCA) is a powerful computational technique that utilizes advanced mathematical concepts to turn a range of correlated features into a minimal set of features called principal components. In PCA, the data stored in a data set is preserved with reduced dimensions depending on an integrated projection of the samples to a subspace created by an orthogonal axle method. Reduced dimensional computational information is chosen in such a way that important data features are specified with hardly any loss of data. In many areas, this reduction is beneficial just as image compression, data representation and so on. Therefore,

PCA is used in a vast area of biomedical image processing applications. For instance, face recognition and image compression, finding patterns in data of high dimension and feature extraction, image segmentation, image registration, image fusion, and de-noising of images. The reduced dimensional space can be calculated by covariance matrix' eigenvectors.

If $x_1 \, x_2 \, \dots \, x_n$ is a set of n $N \times 1$ vectors, $\bar{x}$ is their average:

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iN} \end{bmatrix} \tag{5}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{i=n} \begin{matrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iN} \end{matrix} \tag{6}$$

If X is the $N \times n$ matrix with columns $x_1 - \bar{x}$, $x_2 - \bar{x}$, $\dots \, x_n - \bar{x}$, $\bar{x}$ is the mean value:

$$x = [x_1 - \bar{x} \quad x_2 - \bar{x} \, \dots \quad x_n - \bar{x}] \tag{7}$$

Letting $Q = XX^T$ be the $N \times N$ matrix:

$$Q = xx^T = [x_1 - \bar{x} \; x_2 - \bar{x} \, \dots \, x_n - \bar{x}] \begin{bmatrix} (x_1 - \bar{x})^T \\ (x_2 - \bar{x})^T \\ \vdots \\ (x_n - \bar{x})^T \end{bmatrix} \tag{8}$$

where, Q is the covariance matrix "aka scatter matrix" and in image processing N is often the number of pixels of the image. Each x_j can be written as:

$$x_j = \bar{x} + \sum_{i=1}^{i=n} g_{ji} e_i \tag{9}$$

where, e_i are the n eigenvectors of Q with non-zero eigenvalues. The eigenvectors $e_1 \, e_2 \, \dots \, e_n$ span an eigenspace and $e_1 \, e_2 \, \dots \, e_n$ are $N \times 1$ orthonormal vectors. The scalars g_{ji} are the coordinates of x_j in the space.

$$g_{ji} = (x_j - \bar{x}) \cdot e_i \tag{10}$$

3.4 Data augmentation

When only a small number of training data are present in the dataset, the data augmentation technique plays a significant role in helping the network to learn the needed invariance and robustness characteristics. Data augmentation includes a range of strategies aimed at improving the quantity and efficiency of the data-sets.

Dataset can be expanded through various data augmentation techniques such as translation, horizontal flipping, random crop, rotation, scale transformation, and noise disturbance.

The data samples $D_t : \{x_i\}_{i=1}^M$ at t^{th} iteration are expanded to $\widetilde{D}_t : \{\tilde{x}_i\}_{i=1}^M$ through the previously mentioned data augmentation technique when they are fed into the deep CNN. To express this mathematically, the data augmentation operation of each sample $(x \rightarrow \tilde{x})$ can be viewed as:

$$\tilde{x} = g(x) \tag{11}$$

where, $g(\cdot)$ represents corresponding data augmentation operation. Each different data augmentation technique can be represented by a different $g(\cdot)$.

It assists to avoid over-fitting of the network while also seeking to calibrate the big size of parameters in CNNs. During the suggested model, combinations of patches (or subimages) are generated indiscriminately through obtaining patches from the full images of the used Kaggle dataset. 48x48 dimensional patches are generated by randomly choosing their centers within the complete original retinal image. Patches are selected so that they can slightly located beyond the field of view, helping the neural network to discern the boundary of the field of view and the blood-vessel (vein, artery, or capillary). In spite of that the selected patches interfere such that various patches will comprise similar portions of the source images; so that no more data augmentation is undertaken. This was intended to optimize the model's efficiency while retaining the greatest possible number of features from the source image.

To effectively connect the four parts of the algorithm to form a unified algorithm, follow the following steps:

1. To effectively explain the proposed algorithm let's start by considering the data samples of Kaggle dataset as following: $x \in D : \{x_i\}_{i=1}^{2063}$.
2. Converting data samples into grayscale format and applying a set of transformations as discussed in Eq. (1), we get: $x_{pre} = transformation(x)$.
3. Applying apply edge-preserving guided image filtering to perform contrast enhancement mechanism: $x_{GIF(x_{pre})_p} = \sum_{q \in \omega_k} W_{GIF_{pq}}(G) x_{pre_q}$ where the kernel weights function $W_{GIF_{pq}}(G)$ is defined in Eq. (4).
4. Resizing images' dimension to [256 x 256] and creating [2063 x 256 x 256] data samples.
5. Applying data augmentation to increase the number of samples by generating [48 x 48] dimensional patches by randomly choosing their centers within the complete original retinal image, $x_{aug} = g(x_{GIF})$.
6. Splitting the data samples to train-test sets of 60%-40% respectively.
7. Constructing the proposed CNNs based network of DRUnet.
8. Training DRUnet network using training dataset to perform feature extraction, to get a deep CNN model $M : f(x)$ trained on the data set $D : \{x_i\}_{i=1}^N$.
9. Applying PCA that performs dimension reduction to improve the diagnostic accuracy, the reduced dimensional space can be calculated by covariance matrix' eigenvectors: $x_j = \bar{x} + \sum_{i=1}^{i=n} g_{ji} e_i$ The scalars g_{ji} are the coordinates of x_j in the space: $g_{ji} = (x_j - \bar{x}) \cdot e_i$ where $\bar{x}$ is defined in Eq. (6).
10. Evaluating and testing the trained network $M : f(x)$ using testing dataset $D : \{x_i\}_{i=1}^M$ and classify images using g softmax function: $f(x) = softmax(\cdot)$.

3.5 The proposed network model

The aim of this paper is to introduce a new strategy that brings the strength of convolutional neural networks to the diagnosis of DR. Coupled with using PCA that performs dimension reduction to improve the diagnostic accuracy of diabetes disease [13-15]. PCA is used to make a classifier system more effective. The discriminative features obtained from the proposed CNNs model were utilized. After that, the

most efficient features were determined by using the PCA. Due to the PCA efficient properties it can identify patterns in data of high dimensions and can serve to select a subset of features that preserves as much information present in the complete data as possible. In other words, PCA creates new components that store the most valuable information of the features by capturing a high variance [32]. Recently, several studies have used PCA as a feature extraction technique for classification in health care. Rajagopal et al. [33] compared an automatic classification of cardiac arrhythmia using five different linear and non-linear unsupervised dimensional reduction techniques with the neural network (PNN) classifier. With a minimum of 10 components, fastICA computed an F1 score of 99.83%. Zhang et al. [34] detected breast cancer using an AdaBoost algorithm based on PCA. Negi et al. [35] combined PCA with a feature reduction technique called uncorrelated linear discriminant analysis (ULDA) to obtain the best features that control upper limb motions. Avendano-Valencia et al. [36] applied PCA to time frequency representations (TFR) to reduce heart sounds and improve performance. Kamencay et al. [37] presented a new method using PCA-KNN called the scale-invariant feature transform (SIFT) descriptor in different medical images, which resulted in an accuracy of 83.6% when training 200 images. Ratnasari et al. [38] reduced X-ray images using a threshold-based ROI and PCA. They obtained the best gray-level threshold of 150.

The general proposed model that comprises an efficient convolutional neural network architecture is discussed in this section explicitly. Considering that the retinal images supplied to the model were pre-processed in the first step. Deep supervised learning strategies gain significantly higher efficiency and fulfillment in realistic usage, alongside traditionally supervised and unsupervised strategies. The proposed architecture employs a U-net-influenced neural network with the involvement of residual feedback connections in order to allow the migration and re-use of low-level features. Whilst very deep neural networks training poses a number of difficulties associated with it. Since the gradients may disappear, the forward stream frequently decreases, also the training duration may gruesomely delayed.

In order to overcome such issues, reusing of low level characteristics of shallow layers has been brought to benefit from the prosperity in recent years of deep CNNs. Skip residual links adopt a linkage design which incorporates high level characteristics of dense layers together with low level characteristics of thin layers towards an efficient comprehensive method of classification. Feature vectors from the preceding layers are used for each current layer as entries, as well as its own local features are inserted into the next layers as inputs. There are many inspiring stimuluses for re-use of the low-level characteristics by passing across layers: it mitigates the issue of gradient descent, facilitates the propagation of features, empowers re-use of features and significantly minimizes parameters. Figure 6 provides an illustration of some network middle layer outputs.

Figure 7 (a, b) displays the architecture of the proposed RUnet-PCA deeply CNN. It is guided by the architectural design of U-Network [39]. There are two significant features which concentrate the key contribution of the proposed architectural design [40]; firstly, it offers a highly scalable and customized design of accumulated skip links that cloning the previously trained layers and expanding further identity-mapping channels. Secondly, it optimizes the usage of network computational assets. That is accomplished by an expertly designed architecture, which permits the network's depth and width to be expanded while keeping its computational budget constant. Like the structure of U-Network [39, 40], the architecture of the proposed model comprises a left-side route of contraction and a right-side route of expansion. Both are established by duplicating a unit of building unit which incorporates a collection of transformations with a certain topology. The harmonious introduced architecture leads to a relatively homogeneous, multi-branch network structure with just a few hyper parameters to configure. Whilst prior literature shows empirically that even under strict circumstances of retaining complexity, expanding network cardinality (the number of the series of transformations) can cause a rise in the classification efficiency. Moreover, in situations of rising capacity, expanding cardinality is much more beneficial than deepening or widening [40, 41].

The left-side contraction route is comprised of two parts, the first starts with batch normalization (BN) step pursued directly by rectified linear unit (ReLU), that immediately pursued by convolution process of dimension (3x3). The other portion tends to follow the usual CNN structure accompanied by residual links, this entails the reiterated employment of batch normalization (BN) step pursued directly by rectified linear unit (ReLU), that immediately pursued by convolution process of dimension (3x3). A summation function, which represents the residual connection, follows each convolution operation. The final summation function comprises the result of first section. Both of those basic structures which combines two transformations are pursued by rectified linear unit (ReLU) and max-pooling process of dimension (2x2) and stride (2) to perform downampling. The number of feature vectors is duplicated after each downsampling stage.

Likewise, the right-side expansion route is comprised of upsampling of the features vector by applying unpooling process of dimension (2x2), pursued with concatenation function together with equivalently cropped features vector resulting from the route of contraction [40]. First portion starts with batch normalization (BN) step pursued directly by rectified linear unit (ReLU), that immediately pursued by convolution process of dimension (3x3).

The other portion tends to follow the usual CNN structure accompanied by residual links, this entails the reiterated employment of batch normalization (BN) step pursued directly by rectified linear unit (ReLU), that immediately pursued by convolution process of dimension (3x3). A summation function, which represents the residual connection, follows each convolution operation. The final summation function comprises the result of first section. Both of those basic structures which combines two transformations are pursued by rectified linear unit (ReLU) and unpooling process of dimension (2x2). Cropping is essential as a result of the loss of boundary pixels through every convolution process. The number of feature vectors is halved after each upsampling stage. In the latest stage, convolution process of dimension (1x1) is applied in order to assign the resulting 32-component feature map to its correct class number (that will be two in current situation, where a binary classification task is conducted).

Table 1 obviously refers to sizes of feature vectors in input and output cases within every layer. Figure 7 (a, b) illustrates the residual connections-based deep CNN of the presented model. Moreover, it illustrates the parameters used in the model, such as number of layers and number of filters used in

those layers, and even dimensions of those filters.

Varieties of patches (subimages) of the original pre-processed images were used to confirm the efficiency of the proposed model [40]. 48x48 dimensional patches are generated by randomly choosing their centers within the complete original retinal image. The softmax is mostly used in the classification tasks as the activation function of the output layer, it is responsible for determining the probabilities of each image if related to some specific category. The key feature to softmax activation that it stacks the result likelihoods to be within the scope [0 - 1], much similar the sigmoid activation.

The loss function used in the proposed model is the cross entropy, moreover for optimization the stochastic gradient descent is used [42].

The Rectifier Linear Unit (ReLU) is utilized in the role of activation function subsequent to every convolutional layer and the dropout (0.2) is used among every two successive convolutional layers. Model training was conducted up to 150 epochs and the mini batch size was 32 patches. Figure 8 shows the overall proposed RUnet-PCA model of early stage diabetic retinopathy diagnosis through deep CNNs and principal component analysis.
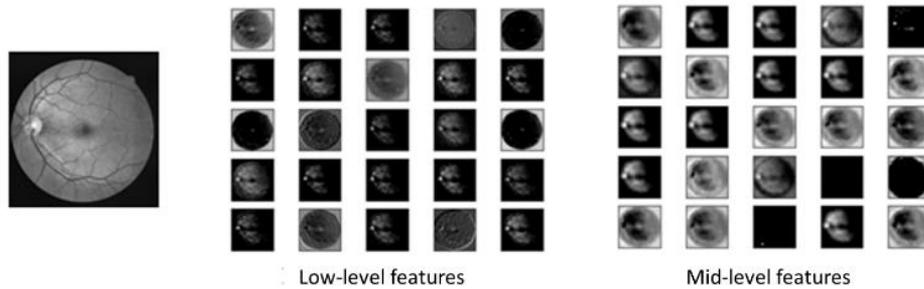
Figure 6. Visualization of some middle layer outputs

(a) Residual U-net structure   (b) ConvDenseBlock structure
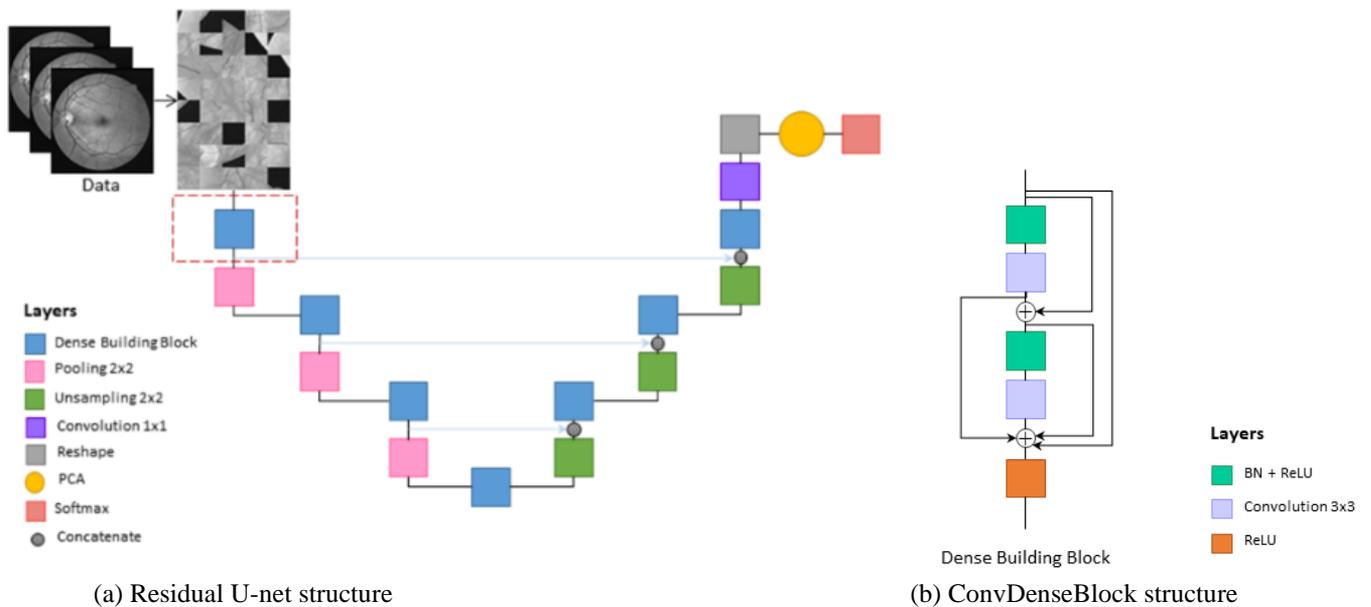
Figure 7. Deep CNN network used in the proposed model
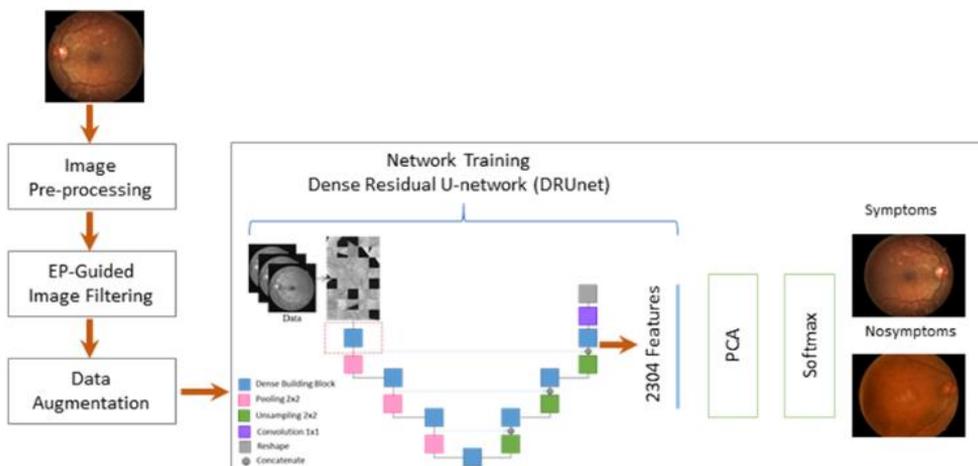
Figure 8. The proposed RUnet- PCA model for diabetic retinopathy diagnosis

Table 1. Size of feature vectors in input / output layers of the proposed model

Layer	Input	Output
Layer 1	(1x48x48)	(32x48x48)
ConvDenseBlock1(32x3x3)		
Layer 2 Pooling1(2x2)	(32x48x48)	(32x24x24)
Layer 3	(32x24x24)	(64x24x24)
ConvDenseBlock2(64x3x3)		
Layer 4 Pooling2(2x2)	(64x24x24)	(64x12x12)
Layer 5	(64x12x12)	(128x12x12)
ConvDenseBlock3(128x3x3)		
Layer 6 Pooling3(2x2)	(128x12x12)	(128x6x6)
Layer 7	(128x6x6)	(256x6x6)
ConvDenseBlock4(256x3x3)		
Layer 8 Unampling1(2x2)	(256x6x6)	(256x12x12)
Layer 9	(384x12x12)	(128x12x12)
ConvDenseBlock5(128x3x3)		
Layer 10 Unampling2(2x2)	(128x12x12)	(128x24x24)
Layer 11	(192x24x24)	(64x24x24)
ConvDenseBlock6(64x3x3)		
Layer 12 Unampling3(2x2)	(64x24x24)	(64x48x48)
Layer 13	(96x48x48)	(32x48x48)
ConvDenseBlock7(326x3x3)		
Layer 14 Conv(2x1x1)	(32x48x48)	(2x48x48)
Layer 15 Reshape	(2x48x48)	(2304x2)

4. RESULTS AND DISCUSSION

This section discusses the experimental data and evaluation metrics, and presents the experimental work and results analysis.

4.1 Data set (Kaggle)

Kaggle diabetic retinopathy dataset, it consists of a large set of high-resolution retina images taken under a variety of imaging conditions and was provided by EyePACS clinics [43]. The image annotation was provided by expert ophthalmologists. Like any real-world data set, it encounters noise in both the images and labels. Images may contain artifacts, be out of focus, underexposed, or overexposed. In this paper, a curated version of the dataset is used [44]. Which is arranged into symptoms and non-symptom, the symptoms set contains 595 images while the non-symptoms set contains 1468 images. Figure 9 shows the class distribution of the Kaggle dataset used in our proposed RUnet- PCA model.

Figure 10 shows ample retina images of Kaggle dataset. The first two images in the top row come from normal subjects, while the two images in the bottom row come from patients who have diabetic retinopathy.
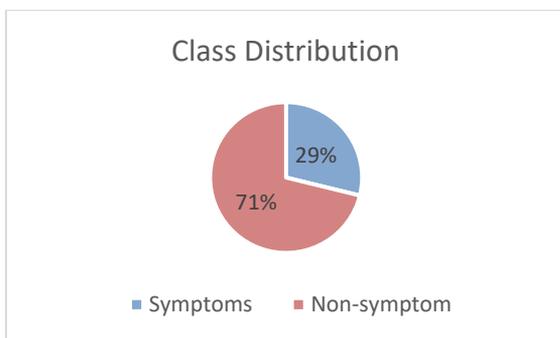
Figure 9. Class distribution of the used Kaggle dataset

Figure 10. Sample retina images of Kaggle dataset. Upper row represents healthy people, lower row represents DR patients

4.2 Experimental work and outputs

To confirm the effectiveness of the proposed model, retinal images of Kaggle diabetic retinopathy dataset are firstly processed and analyzed using some efficient preprocessing methods: grayscale conversion, standardization, contrast-limited adaptive histogram equalization (CLAHE) and gamma adaptation.

Model training is accomplished on a combination of patches (subimages) of the preprocessed full Kaggle images. 48x48 dimensional patches are generated by randomly choosing their centers within the complete original retinal image. Patches are selected that are slightly located beyond the field of view, helping the neural network to discern the boundary of the field of view and the veins, arteries, or capillaries. In spite of the selected patches interfere such that various patches will comprise similar portions of the source images, no more image augmentation is undertaken.

Combinations of differing patches (subimage) numbers are produced by indiscriminately generating a certain number of patches for every image of Kaggle dataset. First 90% of the data augmented is utilized for training, and the remaining 10% is utilized for validation to evaluate the performance of the trained neural network. The loss function used in the proposed model is the cross entropy, moreover for optimization the stochastic gradient descent is used. The Rectifier Linear Unit (ReLU) is utilized in the role of the activation subsequent to every convolutional layer. The softmax function is utilized as an activation for the output layer. Model training was conducted up to 150 epochs.

4.3 Network configuration

This subsection discusses parameter selection and general network configuration, data attributes, training and testing settings as shown in the tables bellow. Table 2 shows data attributes, Table 3 shows training settings.

Table 2. Data attributes

Dimensions of the patches extracted from the full images.	
patch_height	48
patch_width	48

Table 3. Training settings

Number of total patches.	N_subimgs = 15000 (changes)
Number of training epochs.	N_epochs = 150, 200
Number of convolution filters.	Outdim = 32, 64, 128, 256
Size of convolution filters.	(3x3), (1x1)
Size of Max Pooling.	(2x2)
Size of Up Sampling.	(2x2)

The precision of 91.49%, sensitivity of 94.45%, specificity of 98.93% and accuracy of 98.44% were achieved by the proposed RUnet-PCA model. Experimental results show that the proposed RUnet-PCA model achieved a better diagnosis accuracy than other diagnosis methods. Table 4 summarizes the performance analysis of the proposed RUnet- PCA model on Kaggle dataset. Figure 11 displays the loss curve of the proposed model versus epoch.

Table 5 Results of different experiments using the proposed method on the Kaggle database.

Table 5 shows the comparative performance of the proposed model when settings different experiments' parameters. A deep learning model (DRUnet) was used together with a dimensional reduction technique (PCA) to classify the Kaggle diabetes dataset. The key idea of using the dimensional reduction technique is to improve the accuracy of the classifier. A series of experiments using 150 and 200 epochs were performed. In these experiments, the number of subimages fed to the network was changed. Results show that the best performance was obtained when the number of subimages was 17000 and using 150 epochs with classification accuracy of 98.44%, sensitivity of 94.45, and specificity of 98.93, and precision of 91.49. When the number of subimages was 15000 and using 200 epochs the model achieves a classification accuracy of 96.79%, sensitivity of 81.97, and specificity of 98.91, and precision of 91.49. On the other hand, we can notice

that the use of image resolution enhancement techniques in our proposed approach could increase the success of the classification.

In this research, there were two classes as non-symptoms (healthy) and symptoms (patient) which were indicating situation of subject's diabetes disease. Classification results of the proposed system were displayed by using a confusion matrix. In a confusion matrix, each cell contains the raw number of samples classified for the corresponding combination of desired and actual network outputs. The proposed deep learning-based model of and PCA gave very promising results in diagnosing the healthy and patient subjects. The proposed model was arrived to the highest classification accuracy among classifiers in Table 8 It is shown from these results that our proposed method was shown to be a corresponding and safe system to medical diagnostic decision-making.

The efficiency of classification problems can be measured via a confusion matrix, as the relation between classification outcomes and actual outcomes can be observed. Table 6 illustrates a two classes classification problem confusion matrix.
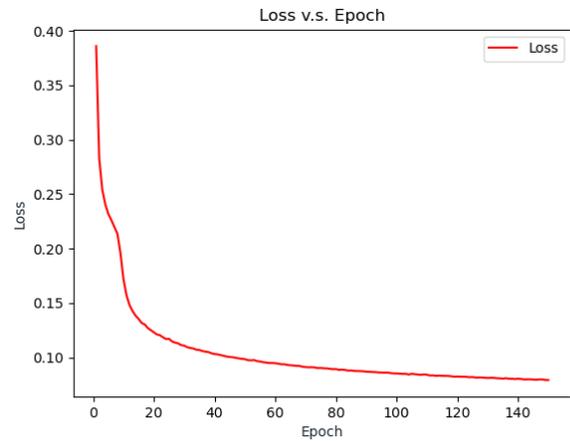
Figure 11. The loss of the proposed RUnet-PCA model against epoch

Table 4. Diagnosis outputs of the proposed RUnet-PCA model on Kaggle dataset

CNN Architecture	Specificity	Sensitivity	Precision	Accuracy
Runet-EGIF (Proposed Model)	98.93%	94.45%	91.49%	98.44%

Table 5. Results of different experiments using the proposed method on the Kaggle database

N epochs	N subimgs	Confusion matrix		ACCURACY	SENSITIVITY	SPECIFICITY	PRECISION
150	10000	3939632	43662	0.952007242	0.686156053	0.98903872	0.89711458
		174136	380713				
	12000	3944663	46731	0.95602827	0.720495145	0.98829206	0.893952494
		152819	393930				
	15000	3896227	40238	0.961794946	0.778715526	0.989778139	0.920911839
		133142	468536				
	16000	3937856	31421	0.954173767	0.689654506	**0.992083949**	**0.925848748**
		176545	392321				
	17000	4002018	43258	**0.984440993**	**0.944506327**	0.989306539	0.914976001
		27351	465516				
200	15000	3927018	43258	0.967914409	0.819762374	0.989104536	0.914976001
		102351	465516				

Table 6. Two classes classification problem confusion matrix

Class	1	2
1	True Positive (TP)	False Negative (FN)
2	False Positive (FP)	True Negative (TN)

— True positive: refers to those correctly classified or detected.
— False positive: refers to those incorrectly classified or detected. It represents the type I error.
— False negative: refers to those incorrectly rejected. It represents the type II error.
— True negative: refers to those correctly rejected.

The efficiency of the proposed model is investigated using well known evaluation measures derived from the confusion matrix, which are accuracy, specify, sensitivity, precision, Jaccard similarity score and error rate.

Accuracy = (TP + TN) / (TP + TN + FP + FN)
Sensitivity = TN / (TN + FP)
Specify = TP / (TP + FN)
Precision = TN / (TN + FN)
Error Rate = (FP + FN / TP + TN + FN + FP)
Jaccard similarity score = 1 - Error Rate

It is recognized that accuracy is the ratio of correctly classified samples to the whole samples, which is a perfect metric to evaluate model's performance. Table 7 shows the confusion matrix of the RUnet-PCA model. Figure 12 displays the receiver operator characteristics (ROC) curve.

Table 7. Confusion matrix of RUnet-PCA model

Class	RUnet-PCA model (Proposed model)	
	Symptoms	Non symptoms
Symptoms	4002018	43258
Non symptoms	27351	465516
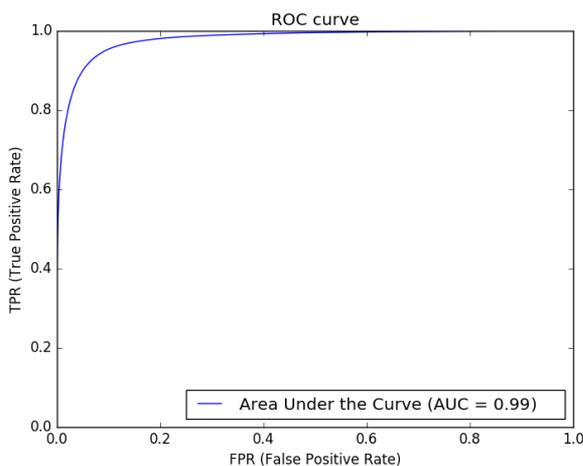
Figure 12. ROC curve of the proposed RUnet-PCA model

The following Table 8 compares the performance of the suggested RUnet-PCA model with other recent deep CNNs architectures that have measured their classification performance in terms of accuracy, sensitivity and specify on Kaggle diabetic retinopathy dataset [45].

As shows in Table 8 the accuracy of VggNet-s model on Kaggle classification is 95.68%. Which is an overall good classification performance. On the other hand, our proposed model achieved a superior classification accuracy of 98.44%. Which proves that our proposed RUnet-PCA model is extremely robust compared to other models and able to achieve effective diagnostic results.

Table 8. Classification outcome of several deep neural network models on Kaggle dataset

Model	Specificity	Sensitivity	Accuracy
AlexNet	94.07%	81.27%	89.75%
VggNet-s	97.43%	86.47%	95.68%
VggNet-16	94.32%	90.78%	93.17%
VggNet-19	96.49%	89.31%	93.73%
GoogleNet	93.45%	77.66%	93.36%
ResNet	95.56%	88.78%	90.40%
Proposed Runet-PCA	98.93%	94.45%	98.44%

5. CONCLUSIONS

The availability of a limited number of ophthalmologists, leads to the idea that an automated system can greatly reduce the tedious manual work involved in diagnosing huge numbers of retinal images. Previous studies have focused in depth and broadly on feature extraction-based diabetic retinopathy diagnosis mechanisms. Then the rapid development of deep convolutional neural networks has become the focus of attention of researchers and the new technology in the classification of retinal images, and in general, in all medical image processing applications [46]. In this paper, we proposed a new early stage diabetic retinopathy diagnosis mechanism which exploits the role of deep convolutional neural networks in diagnosing diabetic retinopathy disease in retinal images. Firstly, we presented a new CNN architecture which is influenced by the well-known U-network with the inclusion of using residual connections to transfer information between layers, which showed superior performance over traditional methods based on feature extraction. Furthermore, principal component analysis (PCA) is used to performs dimension reduction to improve the diagnostic accuracy. As well guided image filtering (GIF) is used to perform as a contrast enhancement for retinal images that not only smooths low gradients, furthermore preserves solid edges. GIF accentuates the edges (traces and lesions in the tissues) of the retina to enable better diagnosis of diabetic retinopathy symptoms. Finally, data augmentation technique is provided, which also improves the performance of the proposed model. The results are encouraging compared to the results of other structures as a tool for diagnosing diabetic retinopathy.

REFERENCES

[1] Klein, R., Klein, B.E., Moss, S.E., Davis, M.D., DeMets, D.L. (1984). The Wisconsin epidemiologic study of diabetic retinopathy. III prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. Archives of Ophthalmology, 102(4): 527-532. https://doi.org/10.1001/archopht.1984.01040030405011

[2] Yau, J.W., Rogers, S.L., Kawasaki, R. (2012). Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care, 35(3): 556-564. https://doi.org/10.2337/dc11-1909

[3] Hiller, R., Sperduto, R.D., Podgor, M.J., Ferris, F.L., Wilson, P.W. (1988). Diabetic retinopathy and cardiovascular disease in type II diabetics. The framingham heart study and the Framingham eye study. American Journal of Epidemiology, 128(2): 402-409. https://doi.org/10.1093/oxfordjournals.aje.a114980

[4] Klein, R., Klein, B.E., Moss, S.E., Cruickshanks, K.J. (1994). The Wisconsin epidemiologic study of diabetic retinopathy. XIV ten-year incidence and progression of diabetic retinopathy. Archives of Ophthalmology, 112(9): 1217-1228. https://doi.org/10.1001/archopht.1994.01090210105023

[5] Klein, R., Klein, B.E., Moss, S.E., Cruickshanks, K.J. (1999). Association of ocular disease and mortality in a diabetic population. Archives of Ophthalmology, 117(11): 1487-1495. https://doi.org/10.1001/archopht.117.11.1487

[6] Venkatesh, P., Sharma, R., Vashist, N., Vohra, R., Garg, S. (2015). Detection of retinal lesions in diabetic retinopathy: Comparative evaluation of 7-field digital color photography versus red-free photography. International Ophthalmolology, 35: 635-640. https://doi.org/10.1007/s10792-012-9620-7

[7] Srivastava, R., Duan, L., Wong, D.W.K., Liu, J., Wong, T.Y. (2017). Detecting retinal microaneurysms and hemorrhages with robustness to the presence of blood vessels. Computer Methods and Programs in Biomedicine, 138: 83-91. https://doi.org/10.1016/j.cmpb.2016.10.017

[8] Marupally, A.G., Vupparaboina, K.K., Peguda, H.K., Richhariya, A., Jana, S., Chhablani, J. (2017). Semi-automated quantification of hard exudates in colour fundus photographs diagnosed with diabetic retinopathy. BMC Ophthalmology, 17: 171-172. https://doi.org/10.1186/s12886-017-0563-7

[9] Larsen, M., Godt, J., Larsen, N., Lund-Andersen, H., Sjølie, A.K., Agardh, E., Kalm, H., Grunkin, M., Owens, D.R. (2003). Automated detection of fundus photographic red lesions in diabetic retinopathy. Investigative Ophthalmology & Visual Science, 44(2): 761-766. https://doi.org/10.1167/iovs.02-0418

[10] Kusakunniran, W., Wu, Q., Ritthipravat, P., Zhang, J. (2018). Hard exudates segmentation based on learned initial seeds and iterative graph cut. Computer Methods and Programs in Biomedicine, 158: 173-183. https://doi.org/10.1016/j.cmpb.2018.02.011

[11] Naqvi, S.A.G., Zafar, H.M.F., Haq, I. (2017). Hard exudates referral system in eye fundus utilizing speeded up robust features. International Journal of Ophthalmology, 10(7): 1171-1174. https://doi.org/10.18240/ijo.2017.07.24

[12] Çinar, A., Yıldırım, M. (2020). Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical Hypotheses, 139: 109684. https://doi.org/10.1016/j.mehy.2020.109684

[13] Velliangiri, S., Alagumuthukrishnan, S., Thankumar Joseph, S.I. (2019). A review of dimensionality reduction techniques for efficient computation. Procedia Computer Science, 165: 104-111. https://doi.org/10.1016/j.procs.2020.01.079

[14] Polat, K., Günes, S. (2007). An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. Digital Signal Processing, 17(4): 702-710. https://doi.org/10.1016/j.dsp.2006.09.005

[15] Shilaskar, S., Ghatol, A. (2013). Dimensionality reduction techniques for improved diagnosis of heart disease. International Journal of Computer Applications, 61(5): 1-8. https://doi.org/10.5120/9921-4538

[16] Krizhevsky, A., Sutskever, I., Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. Proceedings of 26th Annual Conference on Neural Information Processing Systems, pp. 1097-105. https://doi.org/10.1145/3065386

[17] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 770-778. https://doi.org/10.1109/CVPR.2016.90

[18] Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q. (2017). Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 2261-2269. https://doi.org/10.1109/CVPR.2017.243

[19] Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S., Feng, J. (2017). Dual path networks. Proceedings of 30th Annual Conference on Neural Information Processing Systems, pp. 4467-4475.

[20] Veit, A., Wilber, M.J., Belongie, S. (2016). Residual networks behave like ensembles of relatively shallow networks. NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems, pp. 550-558.

[21] Srivastava, R.K., Greff, K., Schmidhuber, J. (2015). Highway networks. CoRR abs/1505.00387. https://arxiv.org/abs/1505.00387.

[22] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2015). Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, pp. 2818-2826. https://doi.org/10.1109/CVPR.2016.308

[23] Zhao, L., Wang, J., Li, X., Tu, Z., Zeng, W. (2017). Deep convolutional neural networks with merge-and-run mappings. CoRR abs/1611.07718. http://arxiv.org/abs/1611.07718.

[24] Xie, S., Girshick, R., Dollár, P., Tu, Z., He, K. (2017). Aggregated residual transformations for deep neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5987-95. https://arxiv.org/abs/1611.05431

[25] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, pp. 1800-1807. https://doi.org/10.1109/CVPR.2017.195

[26] Lin, M., Chen, Q., Yan, S. (2014). Network in network. CoRR abs/1312.4400. https://arxiv.org/abs/1312.4400.

[27] Hu, J., Shen, L., Albanie, S., Sun, G., Wu, E. (2019). Squeeze-and-excitation networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132-41. https://arxiv.org/abs/1709.01507

[28] Zhang, X., Li, Z., Loy, C.C., Lin, D. (2017). PolyNet: a pursuit of structural diversity in very deep networks.

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3900-8. https://arxiv.org/abs/1611.05725

[29] Rahim, S.S., Palade, V., Almakky, I., Holzinger, A. (2019). Detection of diabetic retinopathy and maculopathy in eye fundus images using deep learning and image augmentation. In: Holzinger A., Kieseberg P., Tjoa A., Weippl E. (eds) Machine Learning and Knowledge Extraction. CD-MAKE 2019. Lecture Notes in Computer Science, vol 11713. Springer, Cham. https://doi.org/10.1007/978-3-030-29726-8_8

[30] He, K., Sun, J., Tang, X. (2010). Guided Image Filtering. In: Daniilidis K., Maragos P., Paragios N. (eds) Computer Vision – ECCV 2010. ECCV 2010. Lecture Notes in Computer Science, vol 6311. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15549-9_1

[31] Paris, S., Kornprobst, P., Tumblin, J., Durand, F. (2008). Bilateral filtering: Theory and applications. Foundations and Trends in Computer Graphics and Vision, 4(1): 1-73. https://doi.org/10.1561/0600000020

[32] Guyon, I., Gunn, S., Nikravesh, M. (2008). Feature Extraction: Foundations and Applications. Netherlands: Springer Science and Business Media.

[33] Rajagopal, R., Ranganathan, V, (2017). Evaluation of effect of unsupervised dimensionality reduction techniques on automated arrhythmia classification. Biomedical Signal Processing and Control, 34: 1-8. https://doi.org/10.1016/j.bspc.2016.12.017

[34] Zhang, D., Zou, L., Zhou, X., He, F. (2018). Integrating feature selection and feature extraction methods with deep learning to predict clinical outcome of breast cancer. IEEE Access, 6: 28936-44. https://doi.org/10.1109/access.2018.2837654

[35] Negi, S., Kumar, Y., Mishra, V.M. (2016). Feature extraction and classification for EMG signals using linear discriminant analysis. 2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall), Bareilly, pp. 1-6. https://doi.org/10.1109/icaccaf.2016.7748960

[36] Avendano-Valencia, D., Martinez-Tabares, F., Acosta-Medina, D., Godino-Llorente, I., Castellanos-Dominguez, G. (2009). TFR-based feature extraction using PCA approaches for discrimination of heart murmurs. 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, pp. 5665-5668. https://doi.org/10.1109/iembs.2009.5333772

[37] Kamencay, P., Hudec, R., Benco, M., Zachariasova, M.

(2013). Feature extraction for object recognition using PCA-KNN with application to medical image analysis. 2013 36th International Conference on Telecommunications and Signal Processing (TSP), Rome, pp. 830-834, https://doi.org/10.1109/tsp.2013.6614055

[38] Ratnasari, N.R., Susanto, A., Soesanti, I., Maesadji, (2013). Thoracic X-ray features extraction using thresholding-based ROI template and PCA-based features selection for lung TB classification purposes. In: 2013 3rd international conference on instrumentation, communications. Information Technology and Biomedical Engineering (ICICI-BME), Bandung, pp. 65-69. https://doi.org/10.1109/icici-bme.2013.6698466

[39] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597v1 [cs.CV] 18 May 2015. https://arxiv.org/abs/1505.04597.

[40] Mohammedhasan, M., Uguz, H. (2020). A new deeply convolutional neural network architecture for retinal blood vessels segmentation. International Journal of Pattern Recognition and Artificial Intelligence. https://doi.org/10.1142/S0218001421570019

[41] Lupascu, C.A., Tegolo, D., Trucco, E. (2010). FABC: Retinal vessel segmentation using AdaBoost. IEEE Transactions on Information Technology in Biomedicine, 14(5): 1267-1274. https://doi.org/10.1109/TITB.2010.2052282

[42] Marín, D., Aquino, A., Gegúndez-Arias, M.E., Bravo, J.M. (2011). A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Transactions on Medical Imaging, 30(1): 146-158. https://doi.org/10.1109/TMI.2010.2064333

[43] [dataset] Kaggle Dataset. https://www.kaggle.com/c/diabetic-retinopathy-detection/data, accessed on 8 January 2018.

[44] [dataset] Subset of Kaggle Dataset. https://github.com/javathunderman/retinopathy-dataset

[45] Wan, S., Liang, Y., Zhang, Y. (2018). Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers and Electrical Engineering, 72: 274-282. https://doi.org/10.1016/j.compeleceng.2018.07.042

[46] Neelapu, R., Devi, G.L., Rao, K.S. (2018). Deep learning based conventional neural network architecture for medical image classification. Traitement du Signal, 35(2): 169-182. https://doi.org/10.3166/TS.35.169-182