International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

CNN Architecture for Fundus Image Denoising, Robust Feature Extraction and Classification of Features by Using an Ensemble Classifier in Diabetic Retinopathy

Anand M*[1], Dr. Meenakshi Sundaram A[2]

Submitted: 27/01/2024 **Revised**: 05/03/2024 **Accepted**: 13/03/2024

Abstract: Diabetic retinopathy is considered as the leading cause of vision loss globally. Specifically, it is a microvascular disease that affects the retina resulting in vessel blockage which leads to deprive the retina tissues from nutrition. Effective treatment is possible if it is detected in its early stage because severe and proliferate stages can cause permanent blindness or vision loss.However, this process of DR identification depends on the skills of ophthalmologists, which are sometimes costly and time-consuming. Automatic detection systems were thus developed in an effort to speed up and reduce the cost of the identification process, making it accessible worldwide. The accuracy of the produced predictions, however, was somewhat unsatisfactory for eye doctors to depend on them as diagnostic systems due to the limited credible datasets and medical records for this specific eye illness. Moreover, the captured data suffer from various types of noise. Thus, removing the noise also becomes an important task of this research. As a result, we investigated on a combined denoising and ensemble-based learning technique for filtering and classification. The filtering is done by using CNN based architecture which uses residual noise mapping to identify the noise. In the next stage, we present CNN based feature extraction model and classify the obtained features by using an ensemble classifier. The experimental study shows that the proposed approach achieves classification performance of 97.89, 98.1, and 98.20% in terms of Accuracy, Precision, and Recall for the Kaggle dataset.

Keywords: *Diabetic Retinopathy, Denoising, filering, classification, computer vision.*

1. Introduction

Diabetes mellitus (DM) is a chronic, metabolic, clinically diverse illness whose incidence has been continuously growing worldwide. This health complexity is caused due to decreased ability of the human body to regulate blood sugar. Recently, International Diabetes Federation (IDF) presented an assessment report where it was reported that around 451 million people were affected due to DM in 2017, and is expected to reach 693 million people by 2045 [1]. Diabetes mellitus is defined by chronic hyperglycaemias, which can be caused by decreased insulin production, resistance to insulin's peripheral activities, or both, and finally leads to pancreatic beta-cell failure [2]. Generally, people who are living with DM are more exposed to various other short term and long term complications resulted due to metabolic abnormalities which can lead to severe impact on various organs. These impacts on organs can lead to the development of life-threatening health complications.

These complications are related to microvascular complications such as nephropathy, retinopathy, coronary episode, vision defects, and dementia. Diabetic abnormality is termed based on the body parts which are affected due to abnormal sugar levels such as high blood sugar levels are denoted as diabetic nephropathy which is associated with kidneys, problem related to nerves is known as neuropathy, and problem related to vision is termed as diabetic retinopathy.

Diabetic retinopathy (DR) is the most commonly occurring diabetic complication that arises when retinal blood vessels are damaged due to proliferated sugar levels which lead to vessel enlargement and leakage. Moreover, the advanced stages may lead to complete vision loss. Diabetes used to be an issue that

exclusively affected elderly individuals. But regardless of age, diabetes affects the majority of individuals in today's society, and for those people, it may lead to develop retinopathy-related issues. According to a study presented by WHO [3], in 2014 adults who have 18 years or more have reported diabetes, in 2012 diabetes was identified as direct cause of fatality for 2.2 million people. Similarly, International Diabetes Federation reported that total 463 individuals had diabetes, 88 million cases were from Southeast Asia. Out of 88 million, 77 million cases were reported in India.In 2019, the National Diabetes and Diabetic Retinopathy Survey report estimated the 11.8% occurrence in individuals who are beyond 50 years of age [4]. The prevalence was same for male and female however; it has been noticed as higher in urban territories. According to the survey, the prevalence of diabetic

[1] *Research Scholar, School of Computer Science & Engineering, REVA University, Bengaluru*
[2]*Associate Professor, School of Computer Science & Engineering, REVA University, Bengaluru*
* *Corresponding Author email: r21pcs01@reva.edu.in*

retinopathy was 18.6% in people aged 60 to 69, 18.3% in those aged 70 to 79, and 18.4% in people aged 80 and above. Ages 50 to 59 were assessed to have a reduced prevalence of 14.3% [5].

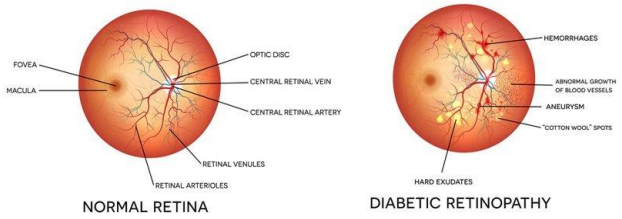


Figure 1. The retinal pictures of healthy and diabetic retinopathy

Diabetic retinopathy is considered as one of the chronic stages of diabetes. As discussed before, the main cause of DR is increased blood sugar levels which can damage the retina's blood vessels [6]. Moreover, veins may constrict, or expand which lead to stop the blood flow. In some instances, the retina may develop new blood vessels. These issues can lead to vision related problems. Figure 1 depicts the sample images of healthy and diabetic retinopathy image. The normal retinal image depicts several parts of retina whereas the DR affected retinal consists of several abnormalities such as haemorrhages, blood vessel growth, cotton wool spots, and aneurysm etc.

Generally, the DR screening is carried out into four stages which are mild non-proliferative, moderate non-proliferative, severe non-proliferative and proliferative retinopathy. The Mild non-proliferative retinopathy is considered as its early stage where microaneurysms (MA) are formed [7, 8]. MAs are characterized by balloon like structures in veins. The next stage is mild non-proliferative retinopathy. In this stage, the veins which support the retina are obstructed. Further, the ailment of retinopathy is termed as severe no-proliferative retinopathy which lead to more obstruction of veins and starts preventing the blood flow in few regions of retina. Finally, the last stage is known as proliferative retinopathy. Here, the growth of a new artery is stimulated by the retina's symptoms of hunger. The arteries are fragile and abnormal. They grow outside of the normal, gel-filled vitreous, which fills the eye, and next to the retina. These veins do not result in symptoms or visual problems on their own due of their thin, sensitive walls. Serious vision loss and a lack of vision may result when such blood vessels leak blood.

This leaking blood appears as spot, known as lesion in the context of retinal images. These lesions are identified as red and bright lesions. The red lesions have MA and HM whereas the bright lesions have soft and hard exudates as presented in figure 1 [9]. The large dark red spots denote the haemorrhages (HM) and smaller red dots are called MA. On the other hand, the hard exudates (EX) appear as bright yellow sport whereas soft exudates (EX) appear as cotton wool and yellowish-white spot which is caused due to damage of nerve fiber. The aforementioned severity level classification of DR can be presented based on lesion classification as presented in below given table 1.

Table 1. Diabetic retinopathy severity level classification based on lesion

Severity level	Lesions
No DR	No lesions
Mild DR	MA Only
Moderate DR	More than MA but less than severe DR
Severe DR	Definite venous beading in more than 2 quadrants, more than 20 intra-retinal HM in each quadrant
Proliferative DR	Pre-retinal HM and neovascularization

1.1. Problem statement

As discussed before, the complications of diabetes, the early detection and prevention of DR is highly recommended to prevent the further complexities. The early detection can be done in the early non-proliferative stage therefore, regular DR screening of diabetic patient through fundus examination is the most effective and widely known method for early detection of DR. However, this process of DR patient screening requires professional clinical knowledge, and experienced ophthalmologists. Generally, the clinicians examine the fundus and retinal image to diagnose the severity of DR but this process consumes more time. Moreover, the professional ophthalmologists are very less to meet the requirement of diagnosis. Several methods have been discussed but these methods still face some challenges such as precise delineation of retinal abnormalities, encompassing microaneurysms, hemorrhages, exudates, and other pathological features, from retinal images remains a formidable challenge. Similarly, the classification of severity also remains challenging issue.

Therefore, biomedical research community has focused on introducing automated process for DR severity detection to improve the efficiency of DR diagnosis. Fundus image analysis has gained attention because these images provide rich information for DR analysis. Several methods have been developed to automate the process of DR detection. Currently, the demand of machine learning based system has increased in biomedical image processing. The main tasks in this field includes image pre-processing, feature extraction, segmentation and classification. The image pre-processing methods play an important role in further stages of DR analysis. The image pre-processing methods include

noise filtering, image rescaling, improving the image quality and many other tasks.

2. Literature survey

This section presents the brief discussion on recent methods related to diabetic retinopathy image processing and classification.

Gayathri et al. [10] reported the advantages of feature extraction and focused on automated classification of DR. Therefore, authors presented feature extraction technique which combines Haralick feature and anisotropic wavelet feature extraction for retinal fundus images. The obtained features were evaluated with the help of different classifiers such as SVM, J48, decision tree etc.

An automated process for DR screening by using computer vision based model was presented by Mookiah et al. [11]. Their work was carried out into multiple stages such as pre-processing, segmentation of optic disc, blood vessels, and exudates. Further, the segmented region was considered for feature extraction where node point count, texture feature and entropy related features were extracted and fed into the probabilistic neural network. In order to enhance the performance of PNN, optimization schemes were also incorporated.

A combined feature extraction technique was presented by Sikder et al. [12],which considers gray level features and texture features from fundus images. Further, ensemble classifier model was developed by using decision tree model. The initial phase performs pre-processing phase steps such as noise removal, contrast enhancement, and image rescaling etc. In the next phase, feature extraction was performed where authors have considered histogram and GLCM based features and concatenate them. Later, a feature selection scheme was also applied. The finally obtained feature set was trained with the help of XGBoost hyper tuning.

Article	Overview
[10]	It used combined feature extract approach by combining Haralick and Wavelet features however this leads to redundancy and require robust mechanism for feature fusion.
[11]	this work performed several steps including pre-processing, segmentation, and employed feature extraction methods (texture, entropy and node point count) and used PNN for classification this model uses basic features to train the model which may result in poor performance for complex image scenarios.
[12]	It performs GLCM and texture feature extraction along with feature selection mechanism. Finally, ensemble classification is used. However, parameter tuning and adding another feature subset can be helpful to achieve better performance.
[13]	This approach uses multilevel feature extraction process and performed ensemble classification but its accuracy remained lower to other approaches due to inefficacy in feature extraction process.
[14]	This model performs severity classification but the complete method is based on the traditional basic feature extraction which may impact the accuracy for large and complex datasets.
[18]	This model used Gaussian matched filter and tree ensemble classifier but. It has considered limited number of features and these features are not normalized to mitigate feature redundancy.
[19]	This model used autoencoder based deep learning approach for DR classification however, in some cases variational model fails to restore images which may lead to poor classification accuracy
[20]	This model mainly focused on only improving the pre-processing stage by introducing Non-Local Means based denoising scheme. the desnoing schemes affect the edge information of biomedical image therefore loss of information affects the segmentation task.
[21]	In this work authors presented DL based CapsNet architecture however the overfitting remains a challenging issue in these architectures.
[22]	In this work, authors used different pretrained models such as ResNet50, ResNet152, and SqueezeNet1. Authors have suggested to train these models on large datasets to obtain the high accuracy.
[23]	This work uses pre-trained DL Models and employs traditional classification models
[26]	This model uses PCNN and CLAHE for feature extraction and then ELM is used for pattern learning. Generally, this model suffer from risk of overfitting and it is sensitive to parameter tuning.
[27]	In [27] pre-trained DL Models for feature extraction and used CNN for classification. Network overfitting and computational complexity for these pre-trained models becomes challenging issue.

Antal et al. [13] presented a computer vision based scheme for automated DR screening. Their method includes several components such as MA detection, exudate detection, macula detection and optic disc detection. Based on these components several features were extracted by considering the pixels and landmarks. The obtained features were fed into an ensemble classifier which uses decision tree, kNN, Naïve Bayes, Random forest, and SVM.

Kandhasamy et al. [14] presented DR severity classification by using multilevel set segmentation approach and SVM classifier. Initially, the morphological clustering operations are performed. The obtained clusters are processed through the local binary pattern feature extraction model to extract the texture features. Moreover, color moments, statistical features such as mean and median are also extracted. The feature selection process is improved by applying the genetic algorithm. Finally, Support vector machine classifier is applied to classify the DR images.

Habib et a. [18] focused on MA detection based on computer vision image processing method. In first stage, it uses Gaussian matched filter to detect the initial candidate set and later a tree based ensemble classifier is presented to classify the extracted features for MA detection.

There is much significance of noise removal in image processing. In a work by Biswas et al. [19], they focused on noise removal before processing the fundus images for classification or any other screening task. In their work, authors adopted deep learning method where deep convolutional denoising auto-encoders were used to remove the noise. The auto encoder uses variation multi-norm loss function. This scheme helps to restore the perceptible details and decrease the noise level.

In another research work, Firdausy et al. [20] discussed that the retinal image segmentation models suffer from accuracy related issues because of poor pre-processing. Therefore, authors suggested to incorporate image pre-processing methods which can reduce the noise and increase the performance. In order to improve the image quality authors presented non-local means filtering method which uses 2D Gaussian model for image modelling. Later, various degrees of smoothing are applied to achieve the denoised image.

Kalyani et al. [21] presented a modified capsule network designed for the detection and classification of diabetic retinopathy. This network includes a convolutional layer and a primary capsule layer for feature extraction. Subsequently, a class capsule layer and a softmax layer are used to estimate the probability of an image belonging to a specific class. The constructed capsule network achieves impressive accuracy rates as 97.98% accuracy for healthy retina images, 97.65% accuracy for stage 1 diabetic retinopathy images, 97.65% accuracy for stage 2 diabetic retinopathy images, and 98.64% accuracy for stage 3 diabetic retinopathy images

Usman et al. [22] reported that the traditional mechanisms ignore the pre-processing task. Therefore, authors addressed this issue by first conducting comprehensive data pre-processing on the Color Fundus Photographs (CFPs). Subsequently, Principal Component Analysis (PCA) is employed for feature extraction. Further, a Deep Learning Multi-Label Feature Extraction and Classification (ML-FEC) model. This model leveraged a pre-trained Convolutional Neural Network (CNN) architecture which includes ResNet50, ResNet152, and SqueezeNet1.

Lahmar et al. [23] conducted an empirical assessment of 28 deep hybrid architectures aimed at automatically classifying referable diabetic retinopathy in binary fashion. In the hybrid architectures, a combination of seven DL techniques for feature extraction, including DenseNet201, VGG16, VGG19, MobileNet_V2, Inception_V3, Inception_ResNet_V2, and ResNet50, was employed. These features were then fed into four different classifiers: Support Vector Machine (SVM), Multilayer Perceptron (MLP), Decision Tree (DT), and K-Nearest Neighbors (KNN).

Gu et al. [24] presented a deep learning based retinopathy classification model which comprised of two main modules as FEB (feature extraction block) and GPB (grading prediction block). This model is composed of two key modules. The FEB plays a primary role in extracting features from fundus images, while the GPB is responsible for categorizing diabetic retinopathy into its five stages. Within the FEB, the transformer exhibits a finer-grained attention mechanism, which enables it to give more focused consideration to regions such as retinal haemorrhages and exudates. Meanwhile, in the GPB, the residual attention mechanism proves effective in capturing the diverse spatial regions associated with various object classes.

Beevi et al. [25] used a unique multilevel severity classification system for Diabetic Retinopathy (DR) is developed through a two-stage process employing SqueezeNet and Deep Convolutional Neural Network (DCNN). Initially, the fundus image is classified into either a normal or abnormal DR category using SqueezeNet. To optimize SqueezeNet, a combination of War Strategy Optimization (WSO) and Fractional Calculus (FC), known as Fractional War Strategy Optimization (FrWSO), is applied. For the abnormal images, the severity level is determined in the second stage using DCNN. Furthermore, the weight adjustment of the DCNN is carried out using the proposed Fractional War Royale Optimization (FrWRO) algorithm, which integrates Battle

Royale Optimization (BRO) with FrWSO. This innovative approach enhances the accuracy and effectiveness of multilevel severity classification for DR.

Nahiduzzaman et al. [26]introduced preprocessing technique known as Contrast Limited Adaptive Histogram Equalization (CLAHE). Subsequently, a parallel convolutional neural network (PCNN) was utilized to extract relevant features. Following this feature extraction step, the classification of DR was carried out using the Extreme Learning Machine (ELM) technique.

Ali et al. [27] developed CNN based approach for DR detection. This model extracts the features using two DL models, Resnet50 and Inceptionv3, and concatenates them before feeding them into the CNN for classification.

3. Proposed model

This section presents the proposed solution for diabetic retinopathy image pre-processing by using computer vision schemes. The proposed approach mainly focusses on image denoising and enhancement to improve the automated image analysis process. The previous section described the important of deep learning scheme in medical image denoising. Therefore, we adopt the CNN based method for image denoising.

3.1. Overview of work

First of all, we perform the pre-processing steps by using proposed pre-processing layer. According to this process, the network considers input I with $n_{ch} \times h \times w$ size and rearranges it into a lower resolution. Further, it extracts the 2x2 patches from reorganized image and rearranges the pixels in different channel as follows:

$$F^0[c,x,y] = I\left[\left\lfloor\frac{c}{3}\right\rfloor, 2x + (c \bmod 2), 2y + \left\lfloor\frac{c}{2}\right\rfloor\right]$$

Further processes are performed on the obtained rescaled image. The complete proposed architecture is depicted in below given figure 2. The complete model consists of attribute extraction model (AEM), enhancement block, selection block and reconstruction block.

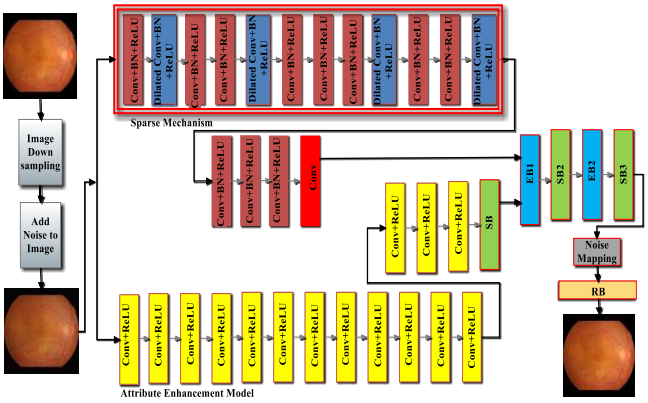


Figure 2. Overall architecture flow

The proposed AEM considers the sparse mechanism to obtain the various attributes. Moreover, it helps to reduce the depth of network. further the obtained features are processed through the enhancement block which improves the feature quality by fusing the attributes. This fusion helps to deal with the real-world noise and various other types of unknown noise. The obtained attributes are processed through the suppression block to distil the features and minimizes the computational cost. Later, reconstruction block is applied to construct the cleaned low resolution image which is further processed through post processing block to generate the final clean image. In this network, the Attribute Extraction Module(AEM) consist of two sub-modules as AEMNet1 and AEMNet2. The first block performs the sparse computation the input is denoted as Y and output is denoted as X. the AEM process can be expressed as:

$$\mathrm{AEM_i} = \mathcal{F}_\mathrm{i}(\mathrm{Y}), \mathrm{i} = 1,2 \ldots \mathrm{n}$$

Where $\mathcal{F}_\mathrm{i}(\mathrm{Y})$ is the feature extraction function and $\mathrm{AEM_i}$represents the obtained features. Similarly, the enhancement block also contains two blocks as EB1 and EB2. The output obtained from $AEM_i$is given to the EB1 to fuse the diverse features from sub-modules of AEM network by following the chain processing of fusing. This process is expressed as:

$$O_{E_1} = E(AEM_1, AEM_2)$$

Where E is the function ofenhancement block EB1 and EB2 and O_{E_1} denotes the outcome of this block.

Similarly, the suppression block has three sub-blocks as CB1, CB2 and CB3. In architecture the CB1 is integrated in AEMNet1, EB and EB2 cover the CB2 which helps to refine the obtained features as follows:

$$O_{CB2} = C_1(O_{E1})$$

Where C_1 is the function of suppression blocks CB1, CB2, and CB3. With the help of the CB2 and EB2 blocks, we obtain the complementary information of attributes which is represented as:

$$O_{E2} = E(O_{CB2}, Y)$$

Similarly, the outcome of suppression block is presented as:

$$O_{CB3} = C_1(O_{E2})$$

Finally, the reconstruction module is used to construct the clean output image from Y with the help of residual features by using the residual operation as:

$$X = Y - O_{CB3}$$

Where $-$ is the residual operation

Further, the obtained filtered image is processed through the CNN based feature extractor module where we extract total 256 features for each image and fed to ensemble classifier to obtain the final classification.

3.2. Attribute Extraction Module for filtering

This model describes the proposed attribute extraction process in proposed deep learning model. As discussed before, this module is divided into two subnetworks which consist of different configurations. The first subblock has a combination of three modules where first module contains convolution, batch normalization and ReLu, second module is constructed with the help of dilated convolution, batch normalization, and ReLu and third module contains convolution block. The combination of these three modules is used to perform he sparse mechanism. The construction of sparse mechanism is depicted in the main architecture d. the convolution layer has a filter size of 3x3,and further, 64x3x3x64 size is allocated to the remaining layers from 2 to 16 whereas the first layer has size of cx3x3x64 where c denotes the channel number. The sparse magnification helps to capture the rich contextual information with the help of dilated convolutions with a dilation factor of 2. Thus, the AEMNet1 can be expressed as:

$$FEB_1 = C\left(CBR_3\left(S(CBR_1(Y))\right)\right)$$

Where CBR denotes the function of Conv+BN+ReLu, S is the sparse mechanism, and C denotes the 3x3 convolution. Similarly, the second sub-network consist of two different modules as Conv+ReLU and CB1 where 3x3 convolution is applied followed by ReLU. Similarly, the CB1 is implemented with 1x1 convolution layer. Similar to the previous block, in this also 2-15 layers are of size 64x3x3x64 whereas the first layer is cx3x3x64 and last layer is designed with 64x1x1x64. The FEBNet procedure can be formulated as:

$$FEB_1 = C_1\left(CR_{15}\left(CR_{15}(Y)\right)\right)$$

CR_{15} is the function of 15th layer which contains Conv+ReLUblocks.

3.3. Attribute Enhancement module for filtering

In this section, we describe the enhancement module to consider the unknown types of noises such as real-world noise. This block is implemented between AEM and CB3 blocks. Moreover, this block is constructed with the help of EB1 and EB2. These blocks consist of three sub-blocks as fusion module, BN and ReLU. The fusion block combines the feature from two different network by applying concatenation. The features obtained from the dilated convolutions and CB1 are different which affects the feature distribution therefore, the BN is applied to eliminate the impact of features. Finally, the ReLU is applied to transform the linear features into nonlinear form. This process is expressed as:

$$O_{E_1} = R\left(B\left(CON(C(CBR_3(S(CBR_1(Y)), CB(CR_{15}(Y)\right)\right)\right)$$

Where B denotes the batch normalization, and R is the activation function of ReLU

3.4. Compression Block (CB) and Reconstruction Block (RB)

This module helps to extract the robust feature and reduce them to minimize the computational complexity. This module has three blocks as CB1, CB2 and CB3. The size of CB1 is 64x1x1x4, size of CB2 128x1x1xc and size of CB3 is 2cx1x1xc. Further, these blocks were implemented by 1x1 convolutions which helps to the feature dimensions and increase the accuracy. Finally, the reconstruction block uses Eq. (7) for reconstruction.

3.5. Loss function

In this process, we adopted mean square error as the loss function to measure the performance of proposed deep learning model for image denoising. This loss function find the relation between predicted image and its corresponding ground truth image $Y_j - X_j$ where X_j represents the j^{th}clean image. Thus, the loss function can be expressed as:

$$L(\theta) = \frac{1}{2N}\sum_{j=1}^{N}\left\|R(Y_j, \theta) - (Y_j - X_j)\right\|_2^2$$

Where θ denotes the trained model parameters, $\{(Y_j, X_j)\}_{j=1}^{N}$ represents the pair of noisy and clean image. Further, it uses Adam optimizer to restore the image.

3.6. CNN based feature extraction for classification

The previous section described the proposed solution for fundus image denoising by applying deep learning based schemes. In this stage, our objective is to apply deep learning scheme for robust feature extraction which can be used for further process. The proposed approach uses convolution layer followed by batch normalization, and max pooling layers. The batch normalization helps to improve the speed of system by rescaling and re-centring the layers. Below given figure 3

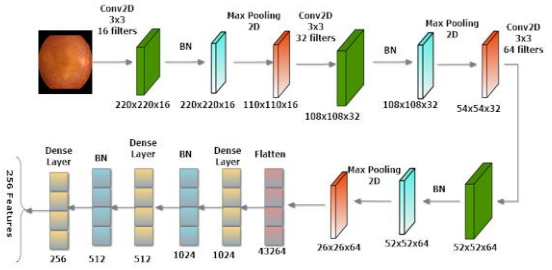


Fig.3.CNN based feature extraction

According to this figure the data is processed through a 2D convolution layer of size 3x3 with filter size of 16, later BN, Max Pooling and 2D convolution operations are performed. However, the size of input filters is varied at each stage as 16, 32 and 64 filters. The max pooling operation helps to capture the largest value from the considered patch, dropout helps to reduce the overfitting and improves the training speed. Further, this model uses Adam optimizer. Finally, the dense layer generates the list of all features. Below given table shows the layer configuration used in this network.

Table 2. Layer configuration

Layer	Shape	Parameters
Input Layer (Conv2D)	[(None, 224, 224,3)]	0
Conv2D	[(None, 220, 220,16)]	1216
Batch Normalization	[(None, 220, 220,16)]	64
Activation	[(None, 220, 220,16)]	0
Max Pooling	[(None, 110, 110,16)]	0
Conv2D_1	(None,108,108,32)	4640
Batch Normalization 2	[(None, 108, 108,32)]	128
Activation	[(None, 108, 108,32)]	0
Max pooling2d	[(None, 54, 54,32)]	0
Conv2D_2	(None,52,52,64)	18496
Batch Normalization 3	(None,52,52,64)	256
Activation_2	(None,52,52,64)	0
Max pooling 2d 2	(None,26,26,64)	0
Flatten	(None, 43264)	0
Dense	(None, 1024)	44303360
Batch normalization 4	(None, 1024)	4096
Activation 3	(None, 1024)	0
Dropout	(None, 1024)	0
Dense 1	(None, 512)	524800
Batch normalization 5	(None, 512)	2048
Activation 4	(None, 512)	0
Dropout	(None, 512)	0
Feature extractor	(None, 256)	131328

3.7. Ensemble classifier

This section presents the proposed solution for ensemble classifier where we consider the CNN features to train the classifiers. In this work, we considerC4.5 random forest and Naïve Bayes classifier. The proposed ensemble model includes bagging and majority vote mechanism to obtain the final outcome.

(a) C4.5 classifier

This classifier is a type of decision tree approach which is developed based on the mechanism of ID3. This approach uses decision tree and traverses through each node of tree to identify the optimal split. The optimal split is obtained based on the maximizing the gain ratio which is expressed as:

$$GainRatio(A) = \frac{Gain\ (A)}{SplitInfo(A)}$$

According to this process, the feature which obtains the highest value of Gain Ratio is considered as splitting attribute. The information gain helps to estimate the uncertainty in the D after partitioning on attribute A. this uncertainty is expressed in the form of entropy as follows:

$$Entropy(D) = -\sum_{x \in X} p(x) \log p(x)$$

Where X denotes the classes in D and $p(x)$ is the proportion of number in classes.

Similarly, the $SplitInfo$ represents the equal splits in attributes and it is presented as:

$$SplitInfo(A) = -\sum_{j=1}^{n} \frac{|D_j|}{|D|} \log\left(\frac{|D_j|}{|D|}\right)$$

Here, $\frac{|D_j|}{|D|}$ represents the weight of j^{th} split

(b) Random forest classifier

This is an another type of decision tree which operates based on the construction of multiple decision tree. It considers multiple inputs and constructs the multiple decision trees to classify the inputs based on the significance of attributes. Generally, it is characterized as ensemble of classification trees where each tree contributes with single vote to achieve the final decision. In this classification, the individual classifier can be presented as:

$$\{h(x, \theta_k)\}, k = 1,2, \ldots i, ..$$

Where h is the random forest classifier, $\{\theta_k\}$ denotes the random vectors which are distributed identically,

(c) Naïve Bayes classifier

This section presents the Naïve Bayes classifier which is based on the probabilistic method to classify the input data. This model uses Bayes theorem, assuming that the existence of one characteristic in a class does not always imply the existence of another feature in the same class. Further, it uses joint probabilities of given classes to estimate the probabilities of given classes.With the help of this assumption, the processing speed of classifier is boosted because it facilitates individual parameter analysis. The final Bayesian network consist of conditional probabilities and structural model. The Bayesian probability of class can be expressed as:

$$P(y|x) = \frac{P(x|y)P(y)}{\sum_{y'} P(x'|y')P(y')}$$

$P(y)$ denotes the probability of occurrence of class y_j , $P(x|y)$ denotes the probability of instance x to be in class y.

4. Result and Discussion

This section presents the outcome of proposed approach for DR screening and classification. The complete approach is implemented by using Python 3.8 running on windows platform. The performance of proposed approach is evaluated for different datasets such as MESSIDOR [15], KAGGLE [16],DIAEETDB0 [17] and DDR dataset.

The MESSIDOR dataset contains total 1200 images. To measure the binary classification performance, 654 images belong to DR category whereas 546 images belong to normal category. For severity grading, the 654 images are divided into three different categories as 153 mild, 247 as moderate and 254 belong to severe DR image.

The Kaggle dataset [16] consist of 35126 images out of which 25810 belong to normal category and 2443, 5292, 873 and 708 images belong to mild NPDR, moderate NPDR, severe NPDR and PDR category, respectively.

Similarly, the DIARETDB0 database contains total 139 images out of which 110 are DR images and 20 are normal images.

The DDR dataset is made publically available by Ocular Disease Intelligent Recognition (ODIR-2019) for segmentation and classification. This dataset comprises of 13673 fundus images which are obtained form 147 different hospitals. In order to perform the classification task, the complete dataset is divided into training, validation and testing sets which consist 6835, 2733 and 4105 images respectively. Moreover, this dataset contaisn six different clasese of DR which are no DR, mild nonproliferative DR, moderate nonproliferative DR, severe nonproliferative DR, proliferative DR, and ungradable.

4.1. Performance measurement parameters

The performance of proposed approach is measured based on confusion matrix calculation. The confusion matrix generated with the help of true positive, false positive, false negative and true negative etc.Below given table 3 shows a sample representation of confusion matrix.

Table 1. Confusion Matrix

Actual class	Predicted class	
	Healthy	DR
Healthy	True Positive	False Negative
DR	False Positive	True Negative

Based on this confusion matrix, we measure several statistical performance parameters such as accuracy, precision, F1-score by using proposed approach.Accuracy is the measurement of correct instance classification out of total number instances. The accuracy is measured as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

Then, we compute Precision of the proposed approach. It is computed by taking the ratio of True Positive and (True and False) positives.

$$P = \frac{TP}{TP + FP}$$

Finally, we compute the F-measure based on the sensitivity and precision values, which is expressed as:

$$F = \frac{2 * P * Sensitivity}{P + Sensitivity}$$

Based on the aforementioned parameters, we consider the weighted average of each class to measure the performance of system. the weighted average performance is measured as:

$$W = \frac{(P_1 * |C_1|) + (P_2 * |C_2|)}{|C_1| + |C_2|}$$

Where $P_1$and P_2 denotes the performance for class C_1 and C_2, respectively.

4.2. Denoising performance

In this section, we describe the image denoising performance and measure the performance in terms of PSNR, RMSE and SSIM. The filtering performance is evaluated for STARE, DRIVE and DIARETDB datasets.

The obtained performance is compared with other methods as presented in below given table.

Table 4. Performance Measurement with various methods

Dataset	Method	PSNR	RMSE	SSIM
DRIVE	CNN-DWT [28]	31.14	0.0095	0.9485
	GAN-CT [29]	31.30	0.0093	0.9535
	RED-CNN [30]	32.25	0.0083	0.9632
	Proposed Model	35.60	0.0062	0.9756
STARE	CNN-DWT [28]	32.40	0.0077	0.9675
	GAN-CT [29]	30.85	0.0094	0.9562
	RED-CNN [30]	31.10	0.0092	0.9381
	Proposed Model	36.50	0.0061	0.9852
DIARETDB	CNN-DWT [28]	31.25	0.0078	0.9710
	GAN-CT [29]	31.28	0.0076	0.9620
	RED-CNN [30]	32.33	0.0094	0.9715
	Proposed Model	35.56	0.0062	0.9865

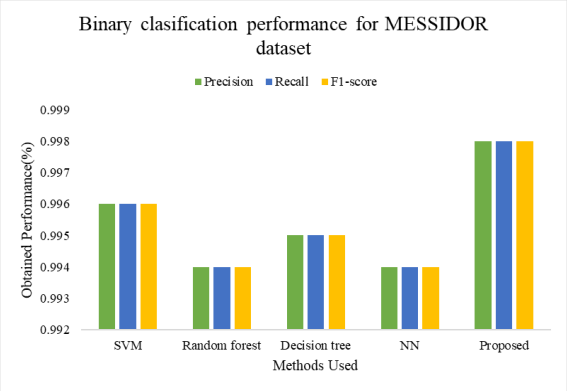
4.3. Classification performance for binary classifier and multiclass scenario

As discussed before, the aforementioned dataset contains images for binary classification also. Therefore, we measure the performance for binary classification and compared the obtained performance with different classifiers. Below given table 5 shows the comparative performance for MESSIDOR Database.

Table 5. Binary classification performance measurement for MESSIDOR database

Classifier	Specificity	Precision	Recall	F1-score
SVM	0.004	0.996	0.996	0.996
Random forest	0.006	0.994	0.994	0.994
Decision tree	0.005	0.995	0.995	0.995
NN	0.006	0.994	0.994	0.994
Inception-V3 [31]	0.9435	0.9512	NA	0.9299
GoogLeNet[31]	0.9375	0.923	NA	0.903
AlexNet[31]	0.9234	0.92	NA	0.9181
ResNet[31]	0.937	0.9415	NA	0.9365
5-Layered CNN [32]	0.9787	0.9815	NA	0.9894
Modified Alexnet [33]	0.9745	92.35	NA	NA
CLAHE+ResNet-101-DELM[34]	0.9819	0.9820	98.20	NA
Proposed	0.989	0.998	0.998	0.998

The traditional InceptionV3 scheme suffer from overfitting and computational complexity leads to increase the training time. Similarly, GoogLeNet also uses Inception modules and it also suffers from vanishing gradient problem which affects the convergence and training stability of the network. The AlexNet is shallower and narrower, which might have limited its ability to capture hierarchical features at various scales.

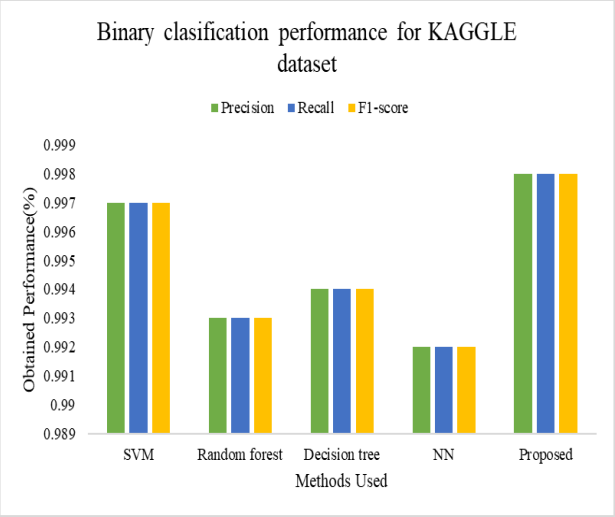


Similarly, we measure the performance of binary classifier for Kaggle dataset and the obtained performance is presented in given table 6.

Table 6. Binary classification performance measurement for KAGGLE database

Classifier	Specificity	Precision	Recall	F1-score
SVM	0.003	0.997	0.997	0.997
Random forest	0.007	0.993	0.993	0.993
Decision tree	0.006	0.994	0.994	0.994
NN	0.008	0.992	0.992	0.992
Proposed	0.002	0.998	0.998	0.998

The comparative study presented for MESSIDOR and KAGGLE shows that the proposed approach achieves better classification performance in terms of Specificity, Precision, Recall, and F1-score.

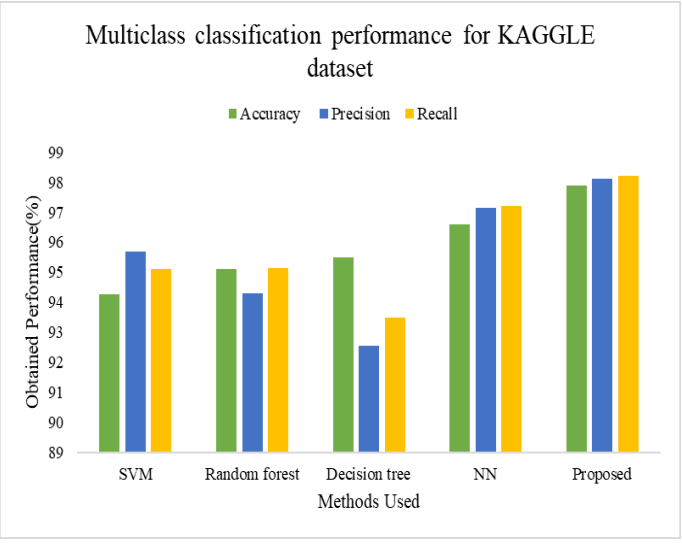


Further, we present the comparative study for multiclass scenario and compared the obtained performance with other methods. Below given table 7 shows the comparative performance in terms of classification accuracy.

Table 7. Multiclass classification performance measurement for KAGGLE database

Classifier	Accuracy	Precision	Recall
SVM	94.25	95.7	95.10
Random forest	95.10	94.30	95.15
Decision tree	95.50	92.55	93.50
NN	96.58	97.15	97.20
Proposed	97.89	98.1	98.20

The complete experimental analysis shows that the proposed approach achieves better performance for image denoising and classification as average classification accuracy is obtained as 99.80% and 97.89% for binary and multiclass classifier.



Similarly, we measure the performance of proposed approach for DDR dataset and compared its performance with existing deep learning based schemes below given table 8 demonstrates the comparative analysis.

Table 8. Multiclass classification performance measurement for DDR database

Model	Class -0	Class -1	Class -2	Class -3	Class -4	Class -5	Avg. Accuracy
VGG-16 [24]	0.9537	0.0423	0.5625	0.3944	0.6436	0.9422	0.5898
ResNet[24]	0.9548	0.0582	0.622	0.3662	0.58180	0.9133	0.5827
GoogLeNet[24]	0.9574	0.0265	0.5759	0.3380	0.5782	0.9162	0.5654
DenseNet[24]	0.8930	0.2275	0.5751	0.4085	0.6364	0.9306	0.6119
SE-BN-Inception[24]	0.9452	0.0476	0.6458	0.1268	0.5818	0.9046	0.5418
ViT+CSRA[24]	0.8280	0.9635	0.7793	0.9881	0.9581	0.9752	0.9154
Proposed Model	0.9623	0.9785	0.8012	0.9915	0.9868	0.9811	0.95.02

The comparative analysis for DDR dataset reported that proposed approach has reported the average accuracy of 93.55% which is better than the state of art DL based methods.

5. Conclusion

This article focus on the development of an automated approach for diabetic retinopathy screening by using computer vision scheme. The complete work is carried out into two phase: first phase focus on image quality enhancement by applying denoising method and second phase focus on feature extraction and developing the ensemble classification model to improve the classification accuracy. The denoising scheme uses CNN based architecture which uses residual image mapping to generate the filtered image. In next phase, we use CNN based model for feature extraction and fed these features into ensemble classifier to generate the final classification outcome. An extensive experimental study is carried out which shows thatproposed approach has obtained the average PSNR as 35.56 dB, 36.50 dB, and 35.60 dB for DIARETDB, STARE and DRIVE datasets/. Similarly, classification study shows that the proposed approach obtained the average classification accuracy as 98.9%, 97.89% and 95.02% for MESSIDOR, KAGGLE and DDR datasets, respectively.

Author contributions

Anand M: Conceptualization, Methodology, Software, Field study, **Anand M:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Dr. Meenakshi Sundaram A:** Visualization, Investigation, Writing-Reviewing and Editing.

Compliance and Ethical Standards

Funding: No funding received for this research work.

***Conflict of interest*: The authors declare there is no conflict of Interest**

Acknowledgement

This endeavour would not have been possible without the generous support from REVA University, Bengaluru. Words cannot express my gratitude to my Research Supervisor Dr. Meenakshi Sundaram A, Associate Professor, School of Computer Science & Engineering, REVA University, Bengaluru and chair of my committee for her invaluable patience and feedback. I also could not have undertaken this journey without my defense committee, who generously provided knowledge and expertise.

REFERENCES

[1] Lin X, Yufeng X, Pan X, Jingya X, Ding Y, Sun Xue, Song Xiaoxiao, Ren Yuezhong, Shan Peng-Fei (2020) Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Sci Rep 10(1):1–11https://doi.org/10.1038/s41598-020-71908-9

[2] Esser, N., Utzschneider, K. M., & Kahn, S. E. (2020). Early beta cell dysfunction vs insulin hypersecretion as the primary event in the pathogenesis of dysglycaemia. Diabetologia, 63(10), 2007-2021.DOI: https://doi.org/10.1007/s00125-020-05245-x

[3] https://www.who.int/news/item/06-04-2016-world-health-day-2016-who-calls-for-global-action-to-halt-rise-in-and-improve-care-for-people-with-diabetes

[4] Kumar, S., Anand, A., Nagarathna, R., Kaur, N., Sivapuram, M. S., Pannu, V., ... & Nagendra, H. R. (2021). Prevalence of prediabetes, and diabetes in Chandigarh and Panchkula region based on glycated haemoglobin and Indian diabetes risk score. Endocrinology, diabetes & metabolism, 4(1), e00162.DOI: https://doi.org/10.1002/edm2.162

[5] Tandon N, Anjana RM, Mohan V, Kaur T, Afshin A, Ong K, Mukhopadhyay S, Thomas N, Bhatia E, Krishnan A, Mathur P (2018) The increasing burden of diabetes and variations among the states of India: the Global Burden of Disease Study 1990–2016. Lancet Glob Health 6(12):1352–1362DOI: https://doi.org/10.1016/S2214-109X(18)30387-5

[6] Shanthi, T., &Sabeenian, R. S. (2019). Modified Alexnet architecture for classification of diabetic retinopathy images. Computers & Electrical Engineering, 76, 56-64.DOI: https://doi.org/10.1016/j.compeleceng.2019.03.004

[7] Sivaprasad, S., & Pearce, E. (2019). The unmet need for better risk stratification of non-proliferative diabetic retinopathy. Diabetic Medicine, 36(4), 424-433.DOI: https://doi.org/10.1111/dme.13868

[8] Putra, R. E., Tjandrasa, H., &Suciati, N. (2020). Severity classification of non-proliferative diabetic retinopathy using convolutional support vector machine. *International Journal of Intelligent Engineering and Systems*, *13*(4), 156-170.DOI: https://doi.org/10.22266/ijies2020.0831.14

[9] Alyoubi, W. L., Shalash, W. M., &Abulkhair, M. F. (2020). Diabetic retinopathy detection through deep learning techniques: A review. *Informatics in Medicine Unlocked*, *20*, 100377.https://doi.org/10.1016/j.imu.2020.100377

[10] Gayathri, S., Krishna, A. K., Gopi, V. P., &Palanisamy, P. (2020). Automated binary and multiclass classification of diabetic retinopathy using haralick and multiresolution features. IEEE Access, 8, 57497-57504.DOI:https://doi.org/10.1016/10.1109/ACCESS.2020.2979753

[11] Mookiah, M. R. K., Acharya, U. R., Martis, R. J., Chua, C. K., Lim, C. M., Ng, E. Y. K., & Laude, A. (2013). Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach. Knowledge-based systems, 39, 9-22.DOI: https://doi.org/10.1016/j.knosys.2012.09.008

[12] Sikder, N., Masud, M., Bairagi, A. K., Arif, A. S. M., Nahid, A. A., &Alhumyani, H. A. (2021). Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. Symmetry, 13(4), 670.DOI: https://doi.org/10.3390/sym13040670

[13] Antal, B., & Hajdu, A. (2014). An ensemble-based system for automatic screening of diabetic retinopathy. Knowledge-based systems, 60, 20-27.DOI: https://doi.org/10.1016/j.knosys.2013.12.023

[14] Kandhasamy, J. P., Balamurali, S., Kadry, S., & Ramasamy, L. K. (2020). Diagnosis of diabetic retinopathy using multi level set segmentation algorithm with feature extraction using svm with selective features. Multimedia Tools and Applications, 79(15), 10581-10596.https://doi.org/10.1007/s11042-019-7485-8

[15] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, ‘‘Feedback on a publicly distributed image database: The Messidor database,’’ Image Anal. Stereol., vol. 33,

no. 3, pp. 231–234, 2014.DOI: https://doi.org/10.5566/ias.1155

[16] Diabetic Retinopathy Detection-Kaggle. Accessed: Oct. 2019. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection/data

[17] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, ''DIARETDB0: Evaluation database and methodology for diabetic retinopathy algorithms,'' Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta Univ. Technol., Lappeenranta, Finland, vol. 73, 2006, pp. 1–17. 1DOI: https://doi.org/10.5244/C.21.15

[18] M. M. Habib, R. A. Welikala, A. Hoppe, C. G. Owen, A. R. Rudnicka, and S. A. Barman, "Detection ofmicroaneurysms in retinal images using an ensemble classifier," Informatics Med. Unlocked, vol. 9, no. May, pp.44–57, 2017.DOI: https://doi.org/10.1016/j.imu.2017.05.006

[19] Biswas, B., Ghosh, S. K., & Ghosh, A. (2020). DVAE: deep variational auto-encoders for denoising retinal fundus image. In Hybrid Machine Intelligence for Medical Image Analysis (pp. 257-273). Springer, Singapore.DOI:https://doi.org/10.1007/978-981-13-8930-6_10

[20] Firdausy, K., Wahyunggoro, O., Nugroho, H. A., Sasongko, M. B., &Hidayat, R. (2019, October). Impact of Different Degree of Smoothing on Non-Local Means based Filter for Retinal Vessel Modeling. In 2019 5th International Conference on Science in Information Technology (ICSITech) (pp. 118-122). IEEE.DOI: https://doi.org/10.1109/ICSITech46713.2019.8987555

[21] Kalyani, G., Janakiramaiah, B., Karuna, A., & Prasad, L. N. (2023). Diabetic retinopathy detection and classification using capsule networks. Complex & Intelligent Systems, 9(3), 2651-2664.DOI: https://doi.org/10.1007/s40747-021-00318-9

[22] Usman, T. M., Saheed, Y. K., Ignace, D., &Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. International Journal of Cognitive Computing in Engineering, 4, 78-88.DOI: https://doi.org/10.1016/j.ijcce.2023.02.002

[23] Lahmar, C., &Idri, A. (2023). Deep hybrid architectures for diabetic retinopathy classification. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11(2), 166-184.DOI: https://doi.org/10.1080/21681163.2022.2060864

[24] Gu, Z., Li, Y., Wang, Z., Kan, J., Shu, J., & Wang, Q. (2023). Classification of diabetic retinopathy severity in fundus images using the vision transformer and residual attention. *Computational Intelligence and Neuroscience*, *2023*.DOI: https://doi.org/10.1155/2023/1305583

[25] Beevi, S. Z. (2023). Multi-Level severity classification for diabetic retinopathy based on hybrid optimization enabled deep learning. Biomedical Signal Processing and Control, 84, 104736.DOI: https://doi.org/10.1016/j.bspc.2023.104736

[26] Nahiduzzaman, M., Islam, M. R., Goni, M. O. F., Anower, M. S., Ahsan, M., Haider, J., & Kowalski, M. (2023). Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier. Expert Systems with Applications, 217, 119557.DOI: https://doi.org/10.1016/j.eswa.2023.119557

[27] Ali, G., Dastgir, A., Iqbal, M. W., Anwar, M., & Faheem, M. (2023). A Hybrid Convolutional Neural Network Model for Automatic Diabetic Retinopathy Classification from Fundus Images. *IEEE Journal of Translational Engineering in Health and Medicine*.DOI: https://doi.org/10.1109/JTEHM.2023.3282104

[28] Kang, E., Min, J., Ye, J.C., et al.: 'A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction', Med. Phys., 2016, 44, (10), pp. e360–e375DOI: https://doi.org/10.1002/mp.12344

[29] Wolterink, J.M., Leiner, T., Viergever, M.A., et al.: 'Generative adversarial networks for noise reduction in low-dose CT', IEEE Trans. Med. Imaging, 2017, 36, (12), pp. 2536–2545. DOI: https://doi.org/10.1109/TMI.2017.2708987

[30] Chen, H., Zhang, Y., Kalra, M.K., et al.: 'Low-dose CT with a residual encoder–decoder convolutional neural network', IEEE Trans. Med. Imaging., 2017, 36, (12), pp. 2524–2535. DOI: https://doi.org/10.1109/TMI.2017.2715284

[31] Bilal, A., Zhu, L., Deng, A., Lu, H., & Wu, N. (2022). AI-based automatic detection and classification of diabetic retinopathy using U-Net and deep learning. Symmetry, 14(7), 1427.DOI: https://doi.org/10.3390/sym14071427

[32] Khan, S. H., Abbas, Z., & Rizvi, S. D. (2019, February). Classification of diabetic retinopathy images based on customised CNN architecture. In 2019 Amity international conference on artificial intelligence (AICAI) (pp. 244-248). IEEE.

[33] Shanthi, T., &Sabeenian, R. S. (2019). Modified Alexnet architecture for classification of diabetic retinopathy images. Computers & Electrical Engineering, 76, 56-64.DOI: https://doi.org/10.1016/j.compeleceng.2019.03.004

[34] Novitasari, D. C. R., Fatmawati, F., Hendradi, R., Rohayani, H., Nariswari, R., Arnita, A., ... &Primadewi, A. (2022). Image fundus classification

system for diabetic retinopathy stage detection using hybrid CNN-DELM. Big Data and Cognitive Computing, 6(4), 146. DOI :https://doi.org/10.3390/bdcc6040146