IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 22, 2021, accepted November 1, 2021, date of publication November 8, 2021,
date of current version November 18, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3125791

Hybrid CNN-SVD Based Prominent Feature Extraction and Selection for Grading Diabetic Retinopathy Using Extreme Learning Machine Algorithm

MD. NAHIDUZZAMAN[1], MD. ROBIUL ISLAM[1], S. M. RIAZUL ISLAM[2], (Member, IEEE), MD. OMAER FARUQ GONI[1], MD. SHAMIM ANOWER[3], (Member, IEEE), AND KYUNG-SUP KWAK[4], (Life Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh
[2]Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea
[3]Department of Electrical and Electronic Engineering, Rajshahi University of Engineering and Technology, Rajshahi 6204, Bangladesh
[4]Department of Information and Communication Engineering, Inha University, Incheon 22212, South Korea

Corresponding authors: Md. Nahiduzzaman (mdnahiduzzaman320@gmail.com) and Kyung-Sup Kwak (kskwak@inha.ac.kr)

This work was supported in part by the National Research Foundation of Korea-Grant funded by the Government of Korea (Ministry of Science and ICT) under Grant NRF-2020R1A2B5B02002478, and in part by Sejong University through the Faculty Research Program under Grant 20212023.

Md. Nahiduzzaman, Md. Robiul Islam, S. M. Riazul Islam, and Md. Omaer Faruq Goni contributed equally to this work.

ABSTRACT This paper exploits the extreme learning machine (ELM) approach to address diabetic retinopathy (DR), a medical condition in which impairment occurs to the retina caused by diabetes. DR, a leading cause of blindness worldwide, is a sort of swelling leakage due to excessive blood sugar in the retina vessels. An early-stage diagnosis is therefore beneficial to prevent diabetes patients from losing their sight. This study introduced a novel method to detect DR for binary class and multiclass classification based on the APTOS-2019 blindness detection and Messidor-2 datasets. First, DR images have been pre-processed using Ben Graham's approach. After that, contrast limited adaptive histogram equalization (CLAHE) has been used to get contrast-enhanced images with lower noise and more distinguishing features. Then a novel hybrid convolutional neural network-singular value decomposition model has been developed to reduce input features for classifiers. Finally, the proposed method uses an ELM algorithm as the classifier that minimizes the training time cost. The experiments focus on accuracy, precision, recall, and F1-score and demonstrate the feasibility of adopting the proposed scheme for DR diagnosis. The method outperforms the existing techniques and shows an optimistic accuracy and recall of 99.73% and 100%, respectively, for binary class. For five stages of DR classification, the proposed model achieved an accuracy of 98.09% and 96.26% for APTOS-2019 and Messidor-2 datasets, respectively, which outperformed the existing state-of-art models.

INDEX TERMS Ben Graham's pre-processing, contrast limited adaptive histogram equalization (CLAHE), convolutional neural network-singular value decomposition (CNN-SVD), diabetic retinopathy (DR), extreme learning machine (ELM).

I. INTRODUCTION

Diabetes mellitus is also known as diabetes which is a collection of metabolic illnesses that happen when a person's blood sugar level is high, and the person does not make enough insulin to control it. The number of the affected person is increasing rapidly day by day. In 2019, International Diabetes Federation (IDF) stated that around 463 million people who have aged between 20 and 79 had been affected by diabetes [1]. Since 2000, the World Health Organization (WHO) has estimated that a 5% increase in diabetes-related early death [2]. Various types of illness such as diabetic retinopathy (DR), stroke, kidney failure, heart attack are caused by diabetes. DR emerges when the retina's blood vessels are disrupted because of excessive blood sugar levels, causes swelling and leakage [3]. In the fundus retina image, this

The associate editor coordinating the review of this manuscript and approving it for publication was Cheng Chin.

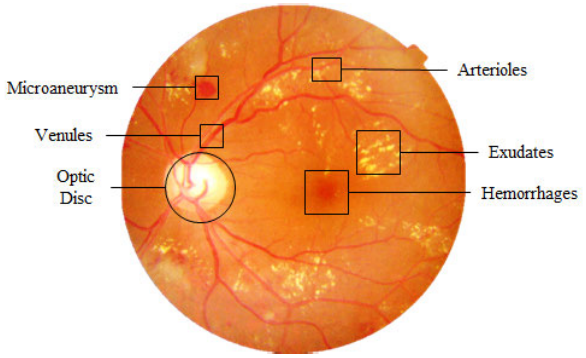

FIGURE 1. Fundus retina images.

leaking blood and fluids show as dots termed lesions. There are two types of lesions: one is red lesions, and another one is right lesions. Figure 1 shows both the lesions where microaneurysms (MA) and haemorrhage (HM) are represented as red lesions on the other hand soft and hard exudates (EX) are represented as right lesions. MA refers to the small dark red dots, while HM refers to the larger spots. Soft EX, often known as cotton wool, looks as yellowish-white and fluffy spots due to nerve fiber injury, whereas hard EX shows as distinctive yellow dots [4]. According to ophthalmologists, DR has been divided into two major stages: non-proliferative DR and proliferative DR (PDR). Non-proliferative DR is further divided into three stages: mild, moderate, severe. Hence, the datasets have five stages: No-DR, mild, moderate, severe, and PDR [5]. The number and types of lesions on the retina image define the stages.

Globally, more than 0.4 million people lost their vision, and around 2.6 million people are affected by severe vision damage [6]. These visual impairments can be prevented or minimized if it is diagnosed and treated promptly enough. But in the early stages of DR, there are few noticeable symptoms; for this reason, most people miss the ideal period for treatment. In the traditional procedure, ophthalmologists use fundus images (FIs) to diagnose DR, which requires a long time, needs a lot of effort, and is vulnerable to misinterpretation of illness. Hence, it is necessary to utilize a computer-aided diagnostic system to detect DR early, prevent misdiagnosis, save money, time, and effort. In the last few years, several deep learning (DL) algorithms have been proposed by researchers for the automatic detection of DR from FIs. In this study, a novel architecture has been proposed for both binary and multiclass DR classification. First, the FIs are pre-processed by Ben Graham's pre-processing method. Further enhanced the contrast of the processed image using CLAHE. Then the prominent features have been extracted using a deep CNN model combined with singular value decomposition (SVD) named the hybrid CNN-SVD method. Finally classified the five stages of DR by using a novel extreme learning machine (ELM) algorithm. These proposed framework performed well in the different datasets, showed an optimistic performance, and surpassed the existing state-of-art models. The key contributions of this study are given below:

- FIs have been pre-processed using BenGraham-CLAHE to reduce the noise, enhance the image contrast, and highlight the lesions.
- This study developed a new hybrid CNN-SVD to extract the features from the FIs. CNN has been used to extract 256 features from processed FIs. Then SVD is applied to reduce these features to 100 features by selecting the most prominent features, decreasing the model complexity, and improving the model performance.
- A comparative analysis has been performed and showed that the proposed method outperformed the state-of-art models in terms of both classification problems in different datasets.

The rest of the paper is organized as: section II represents the recent works on DR classification. Section III describes different steps of the proposed method. The results of the ELM with different approaches are presented in Section IV, and the outcomes compared to the findings of other recent studies. The key conclusions are presented at the end of section V.

II. RELATED WORK

In the last decade, researchers utilized various DL algorithms and many machine learning (ML) algorithms for the automatic detection of DR from FIs. In this section, mainly two categories have been described where one is binary DR classification, and another is multiclass DR classification.

Researchers have used various DL algorithms for binary DR classification from FIs. Das *et al.* utilized maximal principal curvature for extracting the blood vessels from the FIs [7]. Further enhancing and removing the falsely segmented regions by using adaptive histogram equalization (AHE) and morphological opening. Finally, a CNN was used to detect the DR and achieved an accuracy of 98.7%. Liu *et al.* proposed a weighted path CNN, named WP-CNN for the detection of DR [8]. WP-CNN was built by stacking the blocks of the weighted path. WP-CNN was used, including WP-CNN-32, WP-CNN-52, and WP-CNN-105 with 32, 52, and 105 convolutional layers. WP-CNN-105 performed better than the other two models. Pires *et al.* inquisitive a custom CNN for extraction of features from FI for detection of referable DR [9]. By adding data and robust feature-extraction augmentation, the model achieved greater performance. Mahmoud *et al.* proposed a hybrid inductive ML algorithm (HIMLA) for automatic detection of DR from CHASE datasets and achieved an accuracy of 96.62% [10]. Szegedy *et al.* proposed a way to scale up networks in a manner that makes the additional work as efficient as possible by applying appropriately factorized convolutions, and aggressive regularization [11]. Finally, they ensembled 4 models, performed a multi-crop evaluation, and reached a 3.6% top-5 error on the test data.

Gangwar and Ravi developed a hybrid model by merging custom CNN layers on top of the pre-trained

Inception_ResNet-v2 [12] for the detection of five stages of DR [13]. They used both Messidor-1 and APTOS 2019 blindness datasets for training their hybrid model and achieved better performance. Bodapati *et al.* combined a gated-attention mechanism with a deep neural network (DNN) for DR detection from the Kaggle APTOS-2019 dataset [14]. They represented the FI using some pre-trained CNN models. To acquire the reduced versions of these representations without dropping much information, spatial pooling approaches are presented. Their composite model achieved an accuracy of 82.54% for the detection of five stages of DR. Alyoubi *et al.* proposed two deep learning models, which are CNN512 and YOLOv3 [15]. They used CNN512 to detect five DR stages and utilized YOLOv3 for the localization of DR lesions. Finally, they fused these two algorithms for both detection and localize the DR lesions. Dekhil *et al.* utilized transfer learning model VGG16 [16] for the classification of the DR stages from the FI and achieved an accuracy of 77% [17]. Shankar *et al.* developed a hyper-parameter tuning Inception-v4 transfer learning (TL) model for prediction of DR [18]. They utilized CLAHE for enhancing the image quality and performed segmentation using histogram-based segmentation. Majumder *et al.* utilized the EyePACS and APTOS datasets for the detection of DR from FIs [19]. They proposed two deep CNN, which were first trained, one for identifying four stages and the other for further classifying the last stage into two further stages. Qummar *et al.* developed an ensemble model with five TL models, which were Resnet50, Inceptionv3, Xception, Dense121, and Dense169, which enhanced the classification of stages of DR and further encoded the complex characteristics [20].

Sikder *et al.* proposed a decision tree-based ensemble method for the detection of DR [21]. They performed gray-level intensity and utilized a genetic algorithm (GA) to extract the FI features and achieved an accuracy of 94.20%. Gayathri *et al.* extracted the features from FI using a CNN model [22]. They proposed several ML algorithms, for instance, support vector machine (SVM), random forest (RF), multilayer perceptron (MLP), etc. for the automatic detection of DR. Vijayan *et al.* proposed Gabor filter and ML algorithms for the detection of DR and achieved an accuracy of 70.1516% [23]. Tsao *et al.* utilized various ML models, for instance, DT, SVM, artificial neural network (ANN), and logistic regression (LR) for classification of DR [24]. They achieved high accuracy of 79.5% using SVM. Somasundaram and Alli developed an ML bagging ensemble classifier named ML-BEC for the detection of DR from FI [25]. First, they extracted the candidate object from FI and classified the DR using an ensemble classifier. Rajanna *et al.* proposed a neural network (NN) with a combination of pre-processing and augmentation for the classification of DR from FI [26]. Mohammad *et al.* developed an ANN model with three layers for the classification of four stages of DR [27].

Costa *et al.* proposed a multiple instance learning (MIL) framework for classifying DR from FIs [28]. They trained and tested their model using the Messidor dataset and achieved an AUC of 93%. Pour *et al.* enhanced the FIs using CLAHE, and for the classification of DR stages, they used EffiecentNet-5 architecture [29]. They have achieved an AUC of 0.945 for the Messidor-2 dataset. Zeng *et al.* utilized transfer learning, and they developed a unique CNN model with a Siamese-like architecture that can be learned with little effort [30]. They tested their model using 7024 images and achieved a ROC of 0.951. Quellec *et al.* described a general framework for target lesion detection and characterization, which may be quickly done [31]. This tool takes image samples depicting target lesions and other ocular structures similarly shaped but not target and then extracts a feature space from them, allowing a user to more readily locate and choose target lesions in images.

III. PROPOSED DR DIAGNOSTIC FRAMEWORK

Figure 2 demonstrates the proposed framework for the detection of DR. Researchers have been focused on designing a computer-aided system that automatically detects different types of life-threatening diseases from the medical image in the last few decades. Some diseases, such as DR, COVID-19, malaria, etc., require early detection to reduce death rates. Hence it is necessary to develop a system that detects these diseases correctly, efficiently, and on-time to reduce cost and death rates. In this study, a novel framework has been proposed for two scheme detection of DR where one is two class (No DR and DR), and another is five class (No DR, Mild, Moderate, Severe, and Proliferative DR). First, the fundus retina images have been collected. The FIs have been pre-processed using several well-known image processing techniques, for instance, ben graham, CLAHE, etc., which are described later. After pre-processing, a hybrid CNN-SVD model has been developed to extract and select the most discriminant features. Finally, a novel ELM has been proposed for the detection of both binary and multiclass classification.

A. DATASET DESCRIPTION

Two types of the dataset have been used in this study: one comes from the Asia Pacific Tele-Ophthalmology Society (APTOS). The APTOS data had been released as part of the Kaggle blindness detection challenge - 2019 [32]. The retina images were obtained in the lab using several types of clinical cameras, and the database contains 3662 high-resolution colors FIs. Another dataset is Messidor-2 that contains 1748 FIs [33]. Some of the FIs have been incorrectly labelled in this dataset, which are corrected, and the final dataset contains 1738 FIs. There are five stages where no DR, mild, moderate, severe, and proliferative DR are represented by 0, 1, 2, 3, and 4, respectively. The samples of each FI shown in Figure 3. Table 1 shows the data distribution and Table 2 lists the stages and their descriptions.

In this study, two schemes of DR classification have been considered: one is five stages of DR classification, and the other one is binary class classification. For binary class classification stages, 1 to 4 are considered class 1, and 0 is regarded as 0. For both the multiclass and binary class classification,

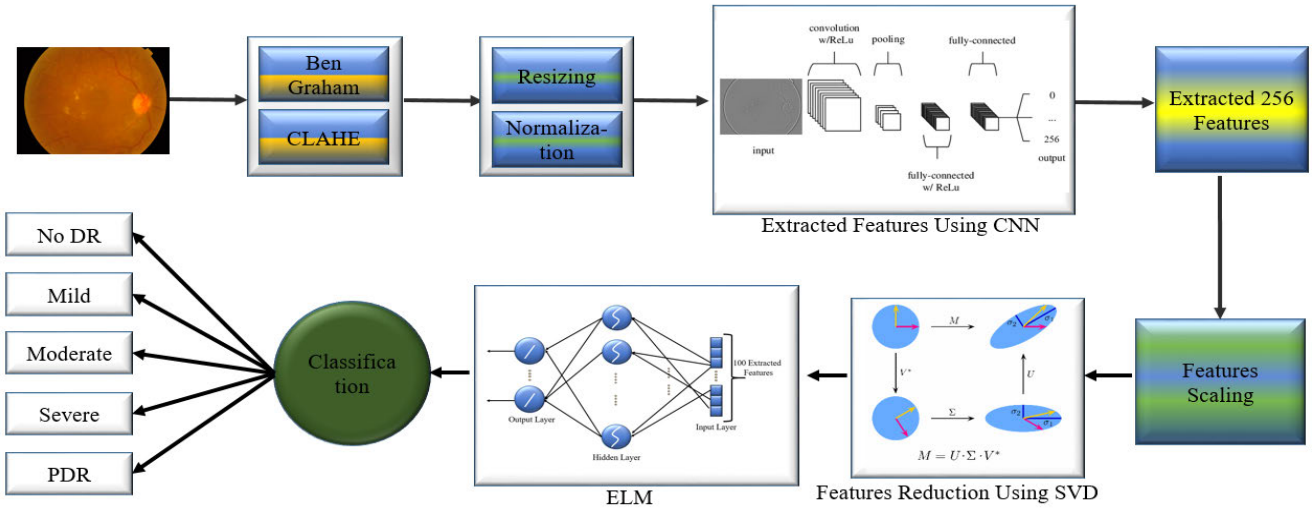

FIGURE 2. Proposed framework for the detection of DR stages.

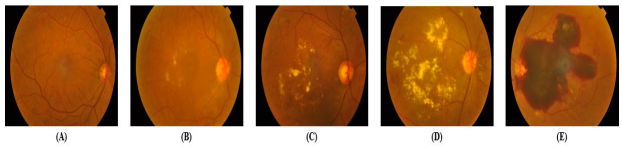

FIGURE 3. Fundus images of (A) No DR, (B) Mild DR, (C) Moderate DR, (D) Severe DR, and (E) PDR.

TABLE 1. Distribution of the dataset.

Stages	APTOS-2019	Messidor-2
No-DR (0)	1805	1017
Mild (1)	370	269
Moderate (2)	999	343
Severe (3)	193	74
Proliferative DR (4)	295	35
Total	3662	1738

the APTOS-2019 has been divided into a ratio of 90:10 for training and testing, respectively. To validate the hyper-parameters of the proposed novel model, another dataset named Messidor-2 has been used for multiclass classification. For training and testing purposes, this dataset is divided into a ratio of 80:20.

B. PRE-PROCESSING

Pre-processing image data is critical because the quality of image pre-processing affects the classification results. In the first pre-processing stages, images have been blurred using the gaussian blur method to reduce the noise. The blurring process has been named Ben Graham as he suggests the parameter used for blurring the image. After blurring the image, CLAHE has been applied to enhance the image contrast. Blurring has been used before CLAHE so that CLAHE does not enhance the unwanted noises in the images. Figure 4 shows the effect of pre-processing. There are five images from five classes. The first column represents the images before pre-processing, the second column represents the blurring images, and the last column represents the contrast-enhanced images.

TABLE 2. The DR stages are defined by the classification of the lesions [34].

Stages	Description of DR Stages
No DR	No spots.
Mild DR	Only microaneurysms (MA).
Moderate DR	There is more to MA but not as severe as DR
Severe DR	**One or more of the following:** In each of the four quadrants, there is more than 20 intraretinal haemorrhage (HM); 2+ quadrants have evident venous beading; 1+ quadrant has prominent intraretinal MA anomalies but no indications of proliferative DR.
Proliferative DR	Neovascularization, pre-retinal HM, or a combination of the two.

In this study, Histogram Equalization (HE) has improved image contrast and model accuracy. It is mainly utilized for low-contrast images associated with scientific activities, such as FIs, X-Ray, Satellite, and Thermal photos [35]. Due to the low contrast of DR images, image contrast enhancement is also helpful for this investigation. This study used a HE approach known as CLAHE. It's a variant of Adaptive Histogram Equalization (AHE). However, in the region where the image is virtually uniform, the AHE overamplifies the noise [36]. The amplification is limited in CLAHE, which overcomes the problem. The clipping factor clips the amplification. A clipping factor of 2 and a tile size of 8×8 have been used in this study. Another thing that has been considered is that after performing Ben graham to FIs, the images are now in the grayscale with a single channel. Since CLAHE works in BGR images, the single-channel has been copied into three channels and performs CLAHE on these processed images, which are shown in Figure 4. From the figure, it is seen that after performing CLAHE, the lessons are more precise.

The datasets include FIs of various forms and sizes converted to a particular shape and size, so that they can be easily fitted into the CNN model. As a result, the FIs are resized

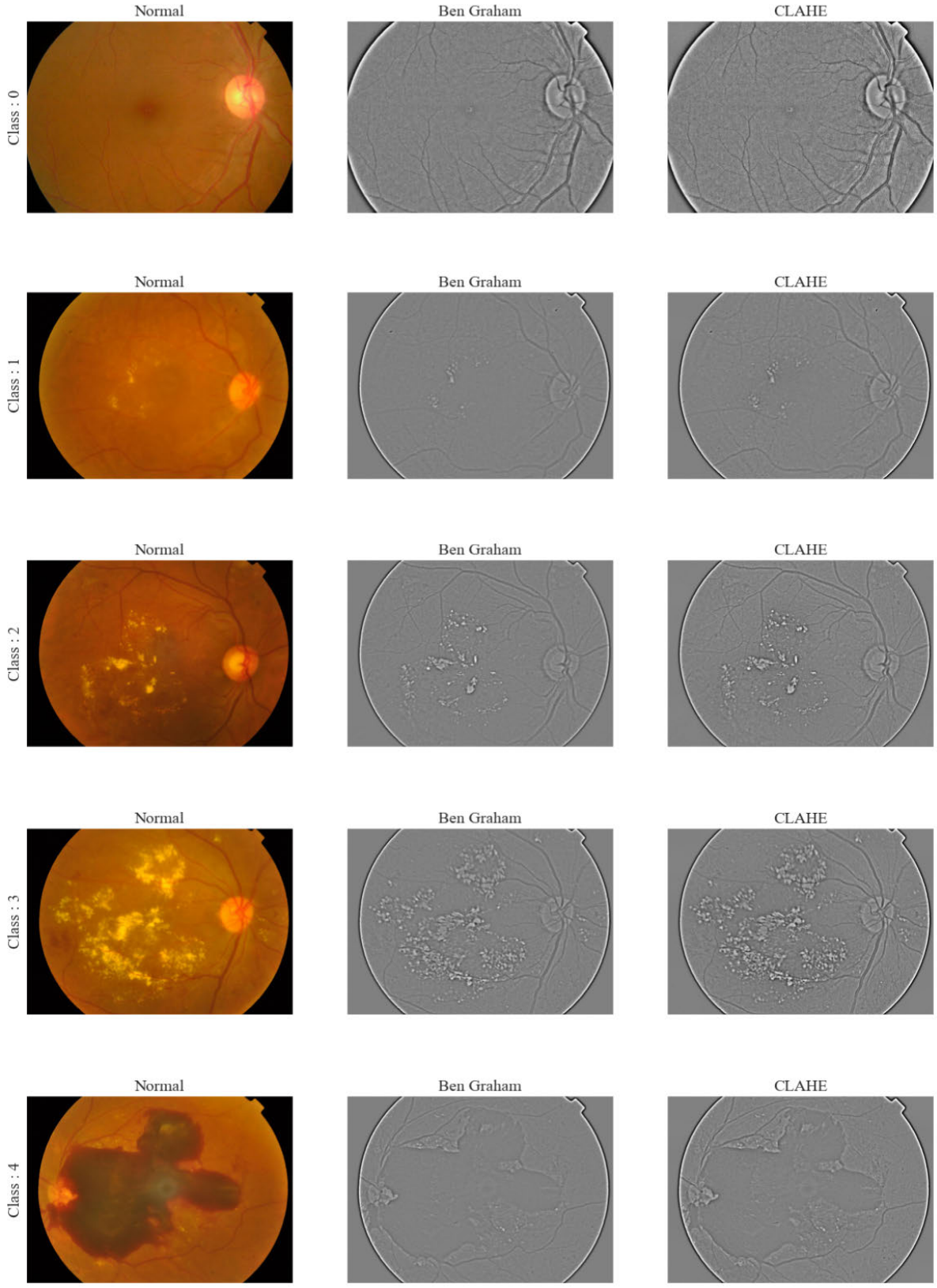

FIGURE 4. Images from different class before and after pre-processing.

to 224 × 224 pixels in size after applying Ben Graham and CLAHE on them. To represent an image, a huge variety of intensity values are used. As a result, normalization is used to prevent unnecessary complexity without using many image

pixels. After reshaping each FI divided by image pixel values by 255; the scale is changed from 0-255 to 0-1, reducing the images' complexity.

Algorithm BenGraham-CLAHE

Input: Fundus image

Output: Pre-processed fundus image

1) **img:** Read image
2) Apply Ben Graham Method:
   a) **blurred_img:** Apply Gaussian Blur on img (Gaussian Kernel Size = (0,0), SigmaX = 10)
   b) **blend_img:** Blending Blurred Image with Original Image (Source1 = img, alpha = 4, source2 = blurred_img, beta = −4, gamma = 128)
3) Create CLAHE: Clip_Limit = 2, Tile Size = (8, 8)
4) **final_img:** Apply CLAHE on blend_img

C. DIMENSIONALITY REDUCTION BASED ON HYBRID CNN-SVD

Sometimes, there are a large number of features available in a dataset where some of them have the least contribution to predict the target variable or create data redundancy. High dimensional feature space has a considerable impact on the performance of a classifier. It is called the curse of dimensionality.

To reduce the model learning complexity and time cost, it is required to apply dimensionality reduction techniques. It transforms the original feature space into a minimal feature space that can hold the actual non-redundant information without significant loss [37]. Several well-known techniques for this purpose include principal component analysis, linear discriminant analysis, and singular value decomposition (SVD). In this study, initially, CNN has been used to extract features. After extracting the features, the features have been standardized. Finally, SVD has been used for dimensionality reduction.

1) FEATURE EXTRACTION USING CNN FROM FUNDUS IMAGES

In this section, a simple CNN has been proposed to extract the most prominent features from the FIs. The model's classification performance will increase if the relevant attributes distinguish between the various DR stages are extracted. For this reason, a simple CNN model has been used. The structure of the feature extractor CNN is illustrated in Figure 5. Table 3 represents the value of the parameter of the CNN model used for this study, and the deep CNN model is summarized in Table 4.

These extracted features can be used for the classification of DR stages efficiently. Each convolutional layer (CL) of CNN has been followed by batch normalization and max-pooling layer. Batch normalization has been employed as it speeds up and improves the performance of the model by re-centering and re-scaling the layers' inputs [38].

TABLE 3. Some parameters value of the CNN feature extractor model.

Parameter Name	Attribute
Max-polling Filter Size	22
Activation Function	ReLu
Dropout	0.5
Loss Function	sparse categorical cross-entropy
Optimizer	Adam
Learning Rate	0.0001
Batch Size	16
Epochs	50

TABLE 4. Summary of proposed simple CNN for feature extraction from fundus images.

Layer (type)	Output Shape	Param
conv2d_input (InputLayer)	[(None, 224, 224, 3)]	0
conv2d (Conv2D)	(None, 220, 220, 16)	1216
batch_normalization(BatchNo1)	(None, 220, 220, 16)	64
activation (Activation)	(None, 220, 220, 16)	0
max_pooling2d (MaxPooling2D)	(None, 110, 110, 16)	0
conv2d_1 (Conv2D)	(None, 108, 108, 32)	4640
batch_normalization (BatchNo2)	(None, 108, 108, 32)	128
activation_1 (Activation)	(None, 108, 108, 32)	0
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 32)	0
conv2d_2 (Conv2D)	(None, 52, 52, 64)	18496
batch_normalization (BatchNo3)	(None, 52, 52, 64)	256
activation_2 (Activation)	(None, 52, 52, 64)	0
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 64)	0
flatten (Flatten)	(None, 43264)	0
dense (Dense)	(None, 1024)	44303360
batch_normalization (BatchNo4)	(None, 1024)	4096
activation_3 (Activation)	(None, 1024)	0
dropout (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 512)	524800
batch_normalization (BatchNo5)	(None, 512)	2048
activation_4 (Activation)	(None, 512)	0
dropout_1 (Dropout)	(None, 512)	0
Feature_Extractor (Dense)	(None, 256)	131328

Max-polling has been used to select the largest value from each cluster's whole neuron and extract the most critical features from the processed FIs [39], [40]. Dropout is employed to reduce overfitting in this case by frequently skipping training all nodes in each layer during the training phase, resulting in a considerable boost in training speed [41]. Adam has been chosen as an optimizer since it performs great when working on large amounts of data [42]. Finally, from each FI, 256 discriminant features have been extracted using the last dense layer.

2) FEATURES REDUCTION USING SINGULAR VALUE DECOMPOSITION (SVD)

It is a data-driven technique based on the generalized concept of fast fourier transform (FFT). From a mathematical point of view, a matrix $D_{(m\times n)}$ can be factored into another three matrices, such as $D = ASB^*$. The factorization will be unique for every matrix. Where, $A_{(m\times m)}$ and $B_{(n\times n)}$ are the unitary matrix. B^* represents the complex conjugate of B. If $MM^* = I$ for matrix M then M is called unitary and for real-valued matrix, $M^* = M^T$. $S_{(m\times n)}$ be the diagonal matrix whose diagonal elements are positively valued in descending order and off-diagonal elements are zero. The number of positive valued diagonal elements is equal to the rank of the matrix D. The rank of a matrix indicates the number of linearly

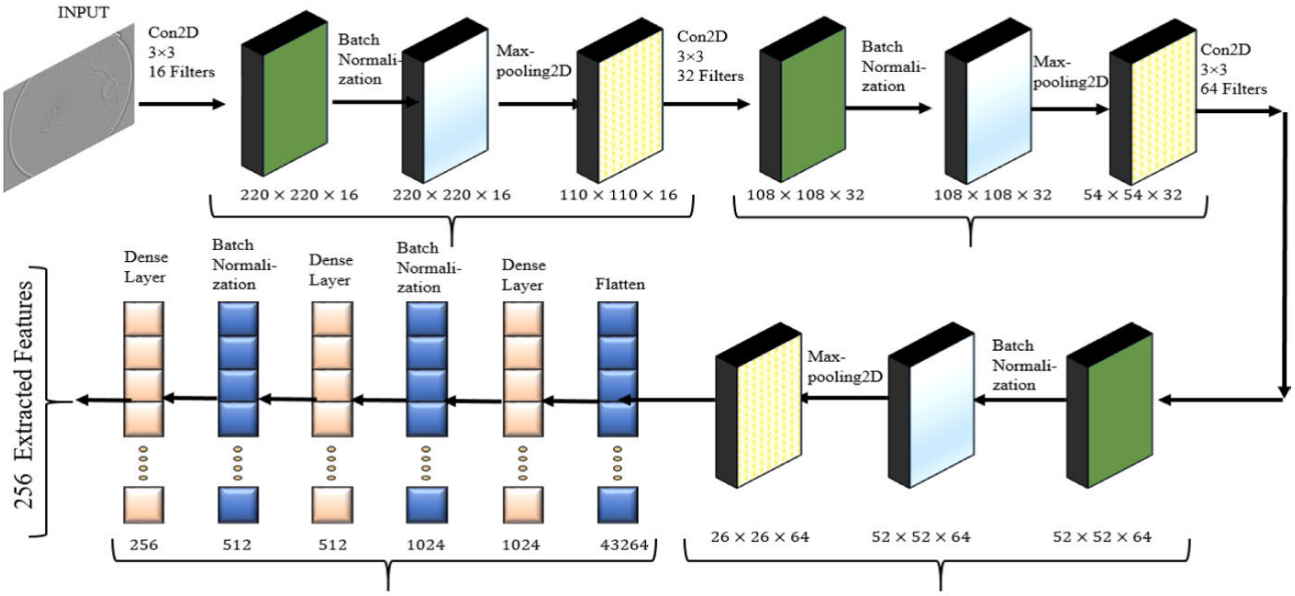

FIGURE 5. Proposed simple CNN for features extraction from fundus images.

independent columns or rows. $D = ASB^*$ can be expressed in series formation as like FFT.

Let, say for example a, s, and b are 3 matrix. Then their multiplication will be

$$\begin{aligned} asb &= \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \\ &= \begin{bmatrix} a_1 s_1 & a_2 s_2 \\ a_3 s_1 & a_4 s_2 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} \\ &= \begin{bmatrix} a_1 s_1 b_1 + a_2 s_2 b_3 & a_1 s_1 b_2 + a_2 s_2 b_4 \\ a_3 s_1 b_1 + a_4 s_2 b_3 & a_3 s_1 b_2 + a_4 s_2 b_4 \end{bmatrix} \\ &= s_1 \begin{bmatrix} a_1 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} + s_2 \begin{bmatrix} a_2 \\ a_4 \end{bmatrix} \begin{bmatrix} b_3 & b_4 \end{bmatrix} \end{aligned}$$

Hence, ASB^* can be written as:

$$\begin{bmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_m \\ | & | & & | \end{bmatrix} \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ | & | & s_m \end{bmatrix} \begin{bmatrix} - & B_1 & - \\ - & B_2 & - \\ & \vdots & \\ - & B_n & - \end{bmatrix}$$

$$= s_1 \begin{bmatrix} | \\ A1 \\ | \end{bmatrix} \begin{bmatrix} - & B1 & - \end{bmatrix} + s_2 \begin{bmatrix} | \\ A1 \\ | \end{bmatrix} \begin{bmatrix} - & B2 & - \end{bmatrix}$$

$$+ \cdots + s_m \begin{bmatrix} | \\ Am \\ | \end{bmatrix} \begin{bmatrix} - & Bn & - \end{bmatrix}$$

SVD can convert D into optimal lower rank approximation by selecting the higher valued elements of S greater than a specified value.

Figure 6 shows the cumulative energy in terms of a number of features. It has been plotted using the normalized cumulative sum of diagonal values of S corresponding to the number of features. It indicates that 100 features have represented approximately 95% of information. Hence, after extracting 256 features using CNN, 100 features have been selected from them.

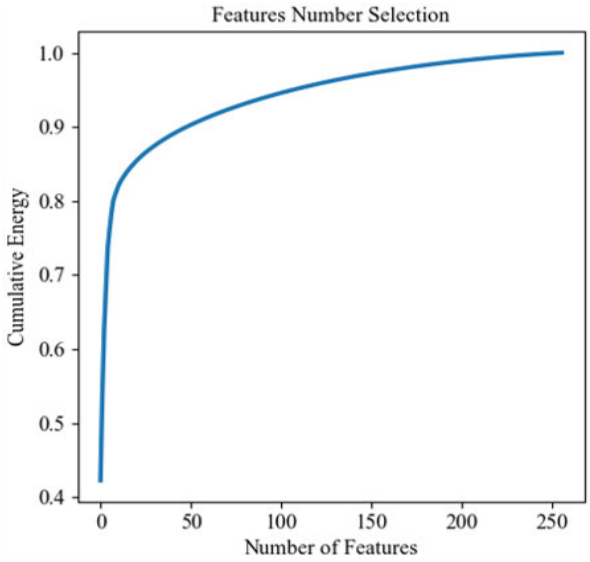

FIGURE 6. Cumulative energy in-terms of number of features.

D. EXTREME LEARNING MACHINE

Huang *et al.* [43] proposed ELM to reduce the training time costing caused by the iterative model parameter tuning method. It's a feed-forwarding NN with an input layer, a single hidden layer, and an output layer. It is capable of achieving the lowest training error. The training process is boosted since the ELM design is straightforward and does not require iterative parameter tuning. The proposed ELM model is shown in Figure 7. It has 100 input nodes, 500 hidden nodes, and

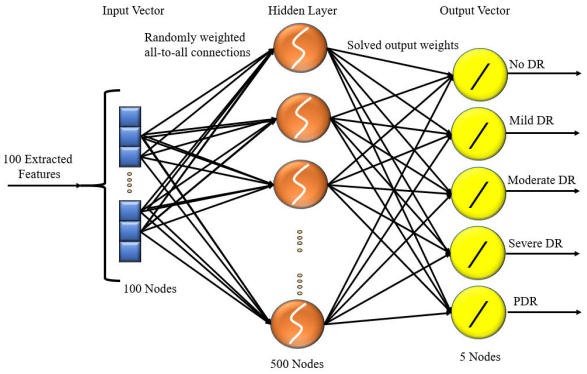

FIGURE 7. Proposed ELM for detection of five stages of DR.

five output nodes for multiclass classification, and for binary classification, there are one output nodes instead of five.

Let's, consider a DR training sample can be represented as $\{X_1, Y_1\} = \{x_{(1,m)}, y_{(1,t)} : m\epsilon R+, t\epsilon R+\}$. X represent the input, and Y represent the output. The train samples can then be expressed using a matrix format such as

$$X_{(n,m)} = \begin{bmatrix} x_{(1,1)} & x_{(1,2)} & \cdots & x_{(1,m)} \\ x_{(2,1)} & x_{(2,2)} & \cdots & x_{(1,m)} \\ x_{(3,1)} & x_{(3,2)} & \cdots & x_{(1,m)} \\ \vdots & \vdots & & \vdots \\ x_{(n,1)} & x_{(n,2)} & \cdots & x_{(n,m)} \end{bmatrix}$$

$$Y_{(n,t)} = \begin{bmatrix} y_{(1,1)} & y_{(1,2)} & \cdots & y_{(1,t)} \\ y_{(2,1)} & y_{(2,2)} & \cdots & y_{(1,t)} \\ y_{(3,1)} & y_{(3,2)} & \cdots & y_{(1,t)} \\ \vdots & \vdots & & \vdots \\ y_{(n,1)} & y_{(n,2)} & \cdots & y_{(n,t)} \end{bmatrix}$$

where, the number of instances is n, the number of attributes is m, and the number of output nodes is t. Hence, for this study $m = 100$ and value of n will be changed with the size of train and test dataset. For binary class $t = 1$ and for multiclass classification $t = 5$. The input weight $W(m, N)$ and bias $B(1, N)$ matrices are then produced at random and N is the number of hidden nodes.

$$W(m,N) = \begin{bmatrix} w_{(1,1)} & w_{(1,2)} & \cdots & w_{(1,N)} \\ w_{(2,1)} & w_{(2,2)} & \cdots & w_{(1,N)} \\ w_{(3,1)} & w_{(3,2)} & \cdots & w_{(1,N)} \\ \vdots & \vdots & & \vdots \\ w_{(m,1)} & w_{(m,2)} & \cdots & w_{(m,N)} \end{bmatrix}$$

$$B(1,N) = \begin{bmatrix} b_{(1,1)} & b_{(1,2)} & \cdots & b_{(1,N)} \end{bmatrix}$$

where, m be the number of attributes and N be the number of hidden nodes. For this study, the value of N will be 500.

Then, calculate the output of the hidden layer using

$$H_{(n,N)} = G(X_{(n,m)} \cdot W_{(m,N)} + B_{(1,N)}).$$

where, $G(x)$ be the activation function. In this study ReLu activation function has been used.

$$H(n,N) = \begin{bmatrix} h_{(1,1)} & h_{(1,2)} & \cdots & h_{(1,N)} \\ h_{(2,1)} & h_{(2,2)} & \cdots & h_{(1,N)} \\ h_{(3,1)} & h_{(3,2)} & \cdots & h_{(1,N)} \\ \vdots & \vdots & & \vdots \\ h_{(n,1)} & h_{(n,2)} & \cdots & h_{(n,N)} \end{bmatrix}$$

Finally, calculate the output layer weight $\beta_{(N,t)}$ by using the Moore-Penrose pseudo inverse.

$$\beta_{(N,t)} = H^{\dagger}_{(N,n)} \cdot T_{(n,t)}$$

T = Output of training data
t = number of outputs

The following steps are used to express the entire procedure:

Algorithm Extreme Learning Machine

1) Randomly generates the input weight $W_{(m,N)}$ and bias $B_{(1,N)}$ matrix.
2) Determine the output $H_{(n,N)}$ of the hidden layer. $H_{(n,N)} = G(X_{(n,m)} \cdot W_{(m,N)} + B_{(1,N)})$
3) Determine the output weight matrix $\beta_{(N,t)}$ $\beta_{(N,t)} = H^{\dagger}_{(N,n)} \cdot T_{(n,t)}$
4) Make prediction using $\beta_{(N,t)}$

IV. EXPERIMENTS AND RESULTS

The machine and deep learning algorithms have been implemented utilizing Keras, with TensorFlow as the backend running on the Pycharm Community Edition19 (2020.2.364) software. A PC with an Intel(R) Core(TM) i7-6700 CPU @3.40GHz, 32GB RAM, and an NVIDIA GeForce GTX1650 SUPER 4 GB GPU running on a 64-bit Windows 10 Pro operating system has been used to train and test the models.

In this study, the model's performance has been measured using an evaluation metric named confusion metrics. The percentage of correctly detected cases among all cases is referred to as accuracy (ACC). It reveals how accurate the classification algorithm is at identifying [44], [45].

$$Acc = \frac{(TP + TN)}{(TP + TN + FP + FN)} * 100\% \quad (1)$$

where TP, TN, FP, and FN indicate true positive, true negative, false positive, and false negative, respectively. The most basic precision (P) is the ratio of TP to all positives forms [46]. In this study, that would be the percentage of patients that accurately identify having a DR out of all the individuals who have it.

$$P = \frac{TP}{(TP + FP)} \quad (2)$$

The recall (R) of a model is the percentage of True Positives it accurately detects [46]. It is a critical factor in this study

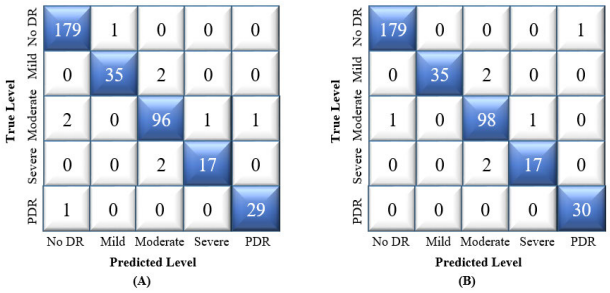

FIGURE 8. CM for detection of five stages (APTOS-2019) of DR (A) with CNN, (B) with hybrid CNN-SVD.

TABLE 5. Results of five stages (APTOS-2019) of DR classification for ELM using only CNN as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No DR	0.98	0.99	0.99	-
Mild DR	0.97	0.95	0.96	-
Moderate DR	0.96	0.96	0.96	-
Severe DR	0.94	0.89	0.92	-
Proliferative DR	0.97	0.97	0.97	-
Average	0.97	0.95	0.96	97.27%

because each DR-affected patient must be identified.

$$R = \frac{TP}{(TP + FN)} \tag{3}$$

The harmonic mean of precision and recall is the F1-score which is expressed mathematically as [47]

$$F1 - score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{4}$$

A. RESULTS FOR FIVE STAGES OF DR CLASSIFICATION

First, the FIs are pre-processed then a simple CNN has been developed to extract 256 features from these processed FIs. Further SVD has been applied for the selection and reduction of these 256 features into 100 features that are more discriminant and finally check whether the performance of ELM has increased or not. Finally, ELM has been used for classification.

1) RESULT FOR APTOS-2019

The proposed ELM has been trained using 3296 data from APTOS-2019 where the numbers of no-DR, mild, moderate, severe, and proliferative DR are 1625, 333, 899, 174, and 265, respectively. To estimate the overall performance of the proposed model, it has been testing using 366 data where the numbers of no-DR, mild, moderate, severe, and proliferative DR are 180, 37, 100, 19, and 30, respectively. A confusion matrix (CM) has been used to examine the ELM model's robustness which is demonstrated in Figure 8.

The precision, recall, f1-score, and accuracy of the ELM have been calculated from this CM. Table 5 represents the result of the proposed ELM model without using SVD. That is to say, the 256 features have been extracted using CNN, then ELM has been applied for the classification of five stages of DR and achieved an accuracy of 97.27%.

TABLE 6. Results of five stages (APTOS-2019) of DR classification for ELM using hybrid CNN-SVD as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No DR	0.99	0.99	0.99	-
Mild DR	1.00	0.95	0.97	-
Moderate DR	0.96	0.98	0.97	-
Severe DR	0.94	0.89	0.92	-
Proliferative DR	0.97	1.00	0.98	-
Average	0.97	0.96	0.97	98.09%

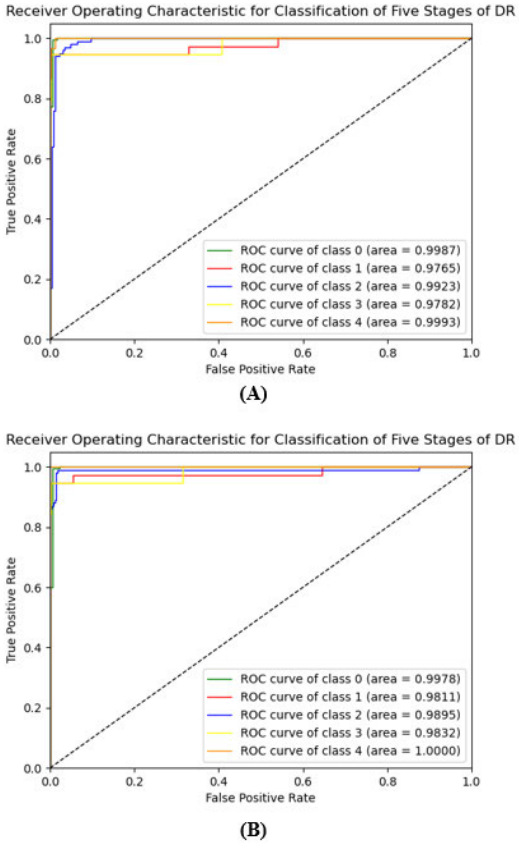

FIGURE 9. ROC for detection of five stages (APTOS-2019) of DR (A) with CNN, (B) with hybrid CNN-SVD.

Further, SVD has been utilized to reduce the features to 100, which also improved the performance of the ELM model with an accuracy of 98.07% demonstrated in Table 6.

The area under the curve (AUC) of the ELM with CNN and hybrid CNN-SVD for the classification of five DR stages is shown in Figure 9. It's a measurement for analyzing a machine learning model's performance and fine-tuning it [48]. Figure 10 represents the graphical comparison of performance based on different metrics between ELM with CNN and ELM with hybrid CNN-SVD for multiclass classification. It can be observed from Figure 10 that processing image by integrating Ben Graham with CLAHE and extracted the features using hybrid CNN-SVD, the ELM model's performance not only increased significantly but also the number of features has been effectively reduced, which decreasing both computational complexity and cost. It also shows that

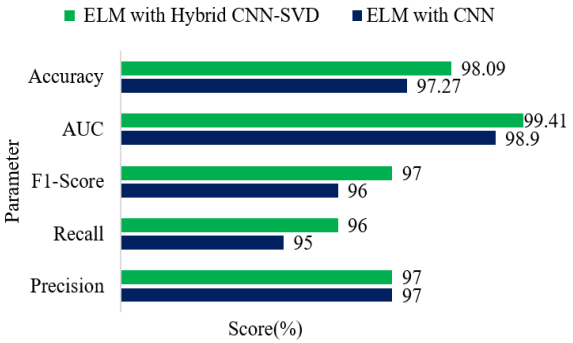

FIGURE 10. Graphical comparison for five stages (APTOS-2019) of DR classification with different approaches.

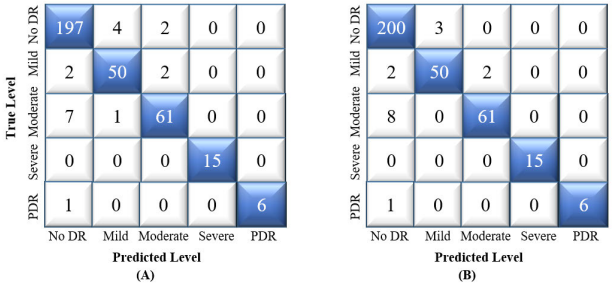

FIGURE 11. CM for detection of five stages (Messidor-2) of DR (A) with CNN, (B) with hybrid CNN-SVD.

TABLE 7. Results of five stages (Messidor-2) of DR classification for ELM using only CNN as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No DR	0.95	0.97	0.96	-
Mild DR	0.91	0.93	0.92	-
Moderate DR	0.94	0.88	0.91	-
Severe DR	1.00	1.00	1.00	-
Proliferative DR	1.00	0.86	0.92	-
Average	0.96	0.93	0.94	94.54%

ELM with hybrid CNN-SVD achieved the highest score in every metric.

2) RESULT FOR MESSIDOR-2

The proposed ELM has been tested using 348 data from the Messidor-2 dataset. The numbers of no-DR, mild, moderate, severe, and proliferative DR are 203, 54, 69, 15, and 7, respectively. Figure 11 shows the CM for both approaches.

The accuracy of the ELM model with CNN-SVD is 96.26%, which is higher than the accuracy of ELM without SVD 94.54%, demonstrated in Tables 7 and 8. The AUC of ELM using CNN-SVD is 98.24%, which is quite good than the ELM using only CNN, which is 97.42%. The ROC of both approaches is demonstrated in Figure 12.

Figure 13 depicts a graphical comparison of performance based on several criteria between ELM with CNN and ELM with hybrid CNN-SVD. Moreover, it demonstrates that ELM with hybrid CNN-SVD achieved a favourable result in every category.

From the performance analysis presented above, the newly structured proposed model can claim its novelty.

TABLE 8. Results of five stages (Messidor-2) of DR classification for ELM using hybrid CNN-SVD as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No DR	0.96	0.99	0.97	-
Mild DR	0.94	0.94	0.94	-
Moderate DR	0.97	0.91	0.94	-
Severe DR	1.00	1.00	1.00	-
Proliferative DR	1.00	0.86	0.92	-
Average	0.98	0.94	0.96	96.26%

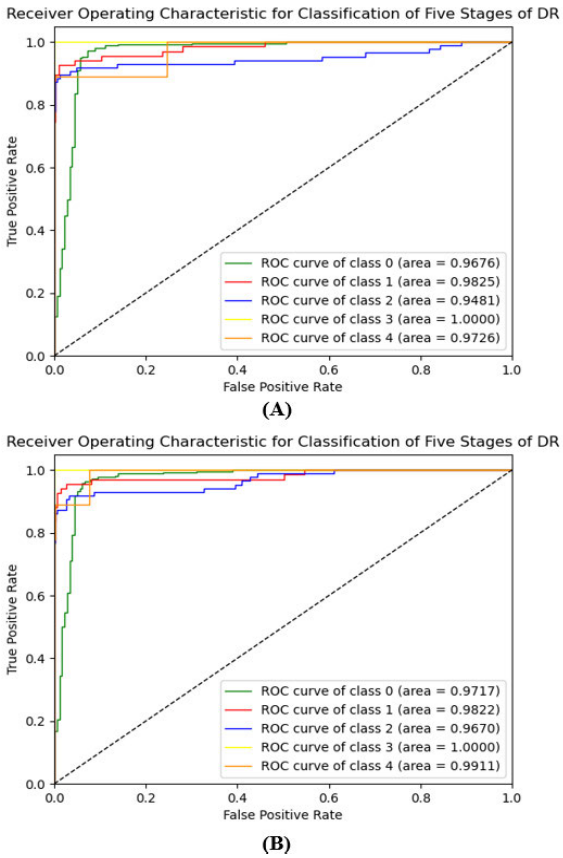

FIGURE 12. ROC for detection of five stages (Messidor-2) of DR (A) with CNN, (B) with hybrid CNN-SVD.

In this study, the combination for data pre-processing (Ben Graham-CLAHE) removed the noise from FIs and enhanced the lesions. Since the lesions are highlighted in the pre-processing stage, the hybrid CNN-SVD extracted the most discriminant features from the lesions accurately and reduced the complexity by removing irrelevant features, which SVD does. Hence, the ELM shows a promising result for both the Aptos-2019 and Messidor-2 datasets with the same hyper-parameter and data pre-processing technique. Therefore, the model hyper-parameter and pre-processing techniques applied in the proposed novel method are independent of the datasets used.

B. RESULTS FOR BINARY DR CLASSIFICATION

For binary class classification of DR similar approach has been taken. The proposed ELM has been trained using 3296 data (APTOS-2019), including 1625 are from

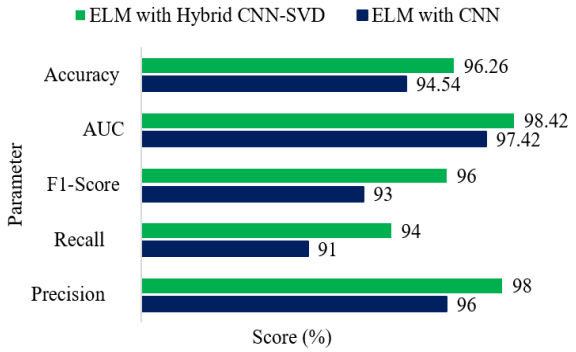

FIGURE 13. Graphical comparison for five stages (Messidor-2) of DR classification with different approaches.

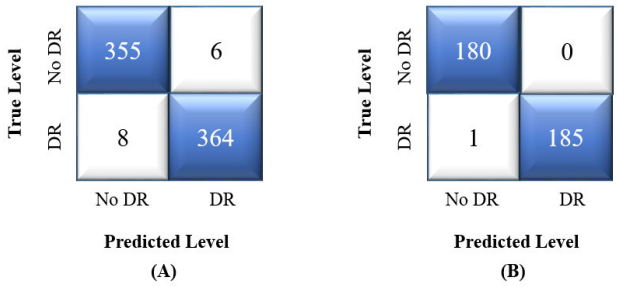

FIGURE 14. CM for binary classification of DR (A) with CNN, (B) with hybrid CNN-SVD.

TABLE 9. Results for binary classification of DR using ELM with CNN as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No-DR	0.98	0.98	0.99	-
DR	0.99	0.98	0.99	-
Average	0.99	0.99	0.99	98.90%

TABLE 10. Results for binary classification of DR using ELM with hybrid CNN-SVD as feature extractor.

Type	Precision	Recall	F1-Score	Accuracy
No-DR	0.99	1.00	1.00	-
DR	1.00	0.99	1.00	-
Average	1.00	1.00	1.00	99.73%

no-DR, and 1671 are from DR. The classification performance of the ELM model has been estimated by testing the model using 366 data where the number of no-DR and DR are 180 and 186, respectively. The CM for binary classification has been shown in Figure 14.

Table 9 shows the result of evaluation metrics of ELM without using SVD. The proposed method achieved accuracy and AUC of 98.91%, and 98.92% respectively.

The performance of the model has been improved using hybrid CNN-SVD and ELM model achieved an accuracy and AUC of 99.73% and 100% respectively, which are shown in Table 10. The ELM model's receiver operating characteristics (ROC) curve for binary classification of DR is depicted in Figure 15.

Figure 16 represents the performance comparison based on different evaluation metrics between ELM with CNN and ELM with hybrid CNN-SVD for binary class classification.

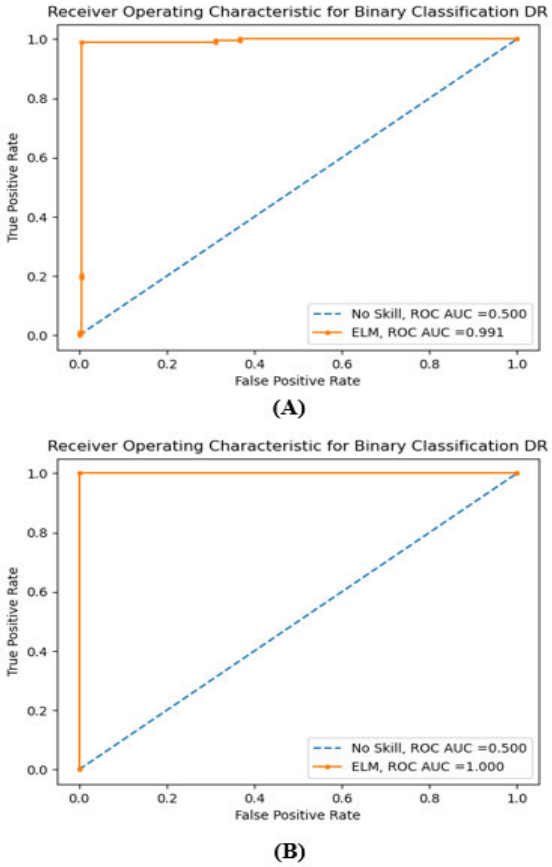

FIGURE 15. ROC for binary classification of DR (A) with CNN, (B) with hybrid CNN-SVD.

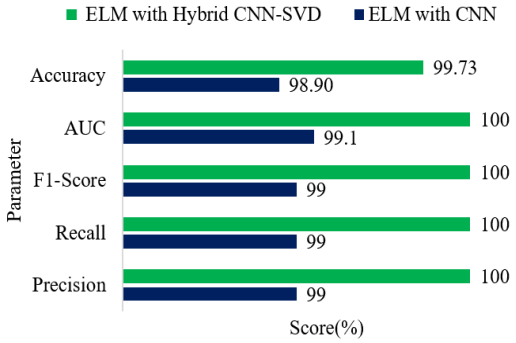

FIGURE 16. Graphical comparison for binary classification with different approaches.

From Figure 16, it is observed that using hybrid CNN-SVD as a feature extractor has been drastically increased the performance of the ELM model with both precision and recall of 100%. In the medical field, recall should be maximized because patients infected with various DR stages must be appropriately diagnosed. It shows that ELM with hybrid CNN-SVD achieved the highest score for each criterion.

C. COMPARISON OF PERFORMANCE TO OTHER WORKS

In this section, the performance of the proposed ELM with hybrid CNN-SVD for DR classification from pro-

TABLE 11. Performance comparison of ELM model for DR classification with other models.

Reference No	No of Class	Dataset	Precision	Recall (Sensitivity)	Accuracy	AUC
[49]	2	APTOS-2019	97.00%	97.00%	97.41%	-
[14]	2	APTOS-2019	98.00%	98.00%	97.82%	98.00%
Proposed Model	2	APTOS-2019	**99.00%**	**99.00%**	**99.32%**	**99.50%**
[50]	5	Messidor-2	-	92.00%	-	88.00%
[51]	5	Messidor-2	-	94.00%	-	90.00%
[52]	5	Messidor-2	-	93.00%	-	94.00%
Proposed Model	5	Messidor-2	**98.00%**	94.00%	**96.26%**	**98.24%**
[49]	5	APTOS-2019	80%	81%	81.7%	-
[53]	5	APTOS-2019	91.37%	-	86.34%	-
[14]	5	APTOS-2019	82.00%	83.00%	82.54%	79%
[54]	5	APTOS-2019	76.00%	77.00%	77.90%	-
Proposed Model	5	APTOS-2019	**97.00%**	**95.00%**	**97.26%**	**98.10%**
[55]	5	APTOS-2019	87.00%	88.24%	83.09%	91.8%
Proposed Model	5	APTOS-2019	**97.00%**	**96.00%**	**97.66%**	**98.96%**
[15]	5	APTOS-2019	89.00%	-	89.00%	97.90%
Proposed Model	5	APTOS-2019	**94.00%**	**94.00%**	**96.31%**	**98.97%**
[21]	5	APTOS-2019	94.34%	92.69%	94.20%	-
Proposed Model	5	APTOS-2019	**95.00%**	**94.00%**	**96.61%**	**98.58%**

cessed FIs has been compared with different existing models. The comparisons have been performed for both binary and five stages of DR classification by utilizing the APTOS - 2019 blindness detection and Messidor-2 datasets. The details of the state-of-art models have been described in section II.

In 2020, Bodapati *et al.* utilized deep CNN for the detection of binary DR classification from APTOS - 2019 FIs [49]. They trained their model using 80% data and 20% data for testing and achieved an accuracy and recall of 97.41% and 97%, respectively. In 2021, Bodapati *et al.* improved the performance of their CNN model using the same data and achieved an accuracy of 97.82% [14]. The same fold has been used for testing the proposed ELM model which achieved an outstanding accuracy and recall of 99.32% and 99%, respectively, which are demonstrated in Table 11.

For multiclass classification, Sanchez *et al.* [50], Seoud *et al.* [51], and Gargeya and Leng [52] utilized the Messidor-2 dataset to test their models and achieved an AUC of 88.00%, 90.00%, and 94.00%, respectively. While using the same dataset, the proposed model achieved an optimistic AUC of 98.24%. Dondeti *et al.* used APTOS-2019 data and achieved an accuracy of 77.90% [54]. Liu *et al.* trained their model using the same data and used 20% data for testing their model and achieved an accuracy of 86.34% [53]. 80% of the data have been used to train the proposed ELM. For calculating the model's performance, it has been tested using 20% data and achieved an accuracy and AUC of 97.26%, and 98.10% respectively. Kassani *et al.* have used 343 data for testing their MLP model and achieved an accuracy and recall of 83.09% and 88.24%, respectively [55]. The same data has been used to test the ELM model and achieved higher accuracy and recall of 97.66% and 96.00%, respectively. Alyoubi *et al.* used 733 data for testing their model and achieved an accuracy of 89% [17]. The same data has been used for testing (733) the proposed model and gained a greater accuracy of 96.31%. Sikder *et al.* have utilized 25% data for testing their model and achieved an accuracy of 94.20% [10]. The ELM model achieved an outstanding accuracy of 96.61% for the same testing data.

From the above comparative study, it can be said the proposed model is an efficient and robust method for early detection of DR. In this study, DR images have been preprocessed using the BenGraham-CLAHE method to improve the image contrast. Hybrid CNN-SVD has been used for dimensionality reduction, where a simple CNN model has been used as a features extractor, and SVD has been used to reduce the number of features. Finally, ELM has been used as the classifier to facilitate the training complexity and time cost. The newly developed framework of this proposed model claims the novelty of this study. The dataset is not so large enough. With a large dataset, the model performance may vary, which is not considered in this study. After all, model performance highly depends on image quality and preprocessing technique applied. DR images that have been used in this study are pretty good in quality. The model performance with low-quality images is out of scope in this study. In the future, the model performance can be analyzed with a large dataset with a mixture of low and high-quality images.

V. CONCLUSION

As diabetes has become more frequent over the world, DR consequences are becoming more common as well. Many DR patients lose their vision globally due to a lack of proper treatment. Early detection and treatment of DR can therefore play a substantial role in reducing the risk of blindness. In this paper, it has been found that the capability of the hybrid CNN-SVD based method to extract features is helpful for binary classification and multiclass classification of DR. The DR classification is more effective when integrated with an ELM approach, shown by the proposed technique. The method has exploited BenGraham's principle and CLAHE image preprocessing methods to reduce the noises, highlight the lesions, and, therefore, achieve a better DR classification performance. According to the comparison with the existing

schemes, it has been concluded that an ELM classifier can detect DR more precisely. The results of this research are expected to be useful for doctors to detect different stages of DR early and diagnose the patients accordingly.

REFERENCES

[1] (2020). *International Diabetes Federation—Facts & Figures*. Accessed: Sep. 10, 2021. [Online]. Available: https://www.idf.org/aboutdiabetes/what-is-diabetes/facts-figures.html

[2] (2021). *Diabetes*. Accessed: Sep. 10, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes

[3] K. Boyd. (2020). *American Academy of Ophthalmology-What is Diabetic Retinopathy*. Accessed: Sep. 10, 2021. [Online]. Available: https://www.aao.org/eye-health/diseases/what-is-diabetic-retinopathy

[4] R. Taylor and D. Batey, *Handbook of Retinal Screening in Diabetes: Diagnosis and Management*. Hoboken, NJ, USA: Wiley, 2012.

[5] J. Nayak, P. S. Bhat, R. Acharya, U. C. M. Lim, and M. Kagathi, "Automated identification of diabetic retinopathy stages using digital fundus images," *J. Med. Syst.*, vol. 32, no. 2, pp. 107–115, 2008.

[6] R. R. A. Bourne, G. A. Stevens, R. A. White, J. L. Smith, S. R. Flaxman, H. Price, J. B. Jonas, J. Keeffe, J. Leasher, K. Naidoo, K. Pesudovs, S. Resnikoff, and H. R. Taylor, "Causes of vision loss worldwide, 1990–2010: A systematic analysis," *Lancet Global Health*, vol. 1, no. 6, pp. e339–e349, Dec. 2013.

[7] S. Das, K. Kharbanda, S. M, R. Raman, and E. D. D, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102600.

[8] Y.-P. Liu, Z. Li, C. Xu, J. Li, and R. Liang, "Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101694.

[9] R. Pires, S. Avila, J. Wainer, E. Valle, M. D. Abramoff, and A. Rocha, "A data-driven approach to referable diabetic retinopathy detection," *Artif. Intell. Med.*, vol. 96, pp. 93–106, May 2019.

[10] M. H. Mahmoud, S. Alamery, H. Fouad, A. Altinawi, and A. E. Youssef, "An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm," *Pers. Ubiquitous Comput.*, vol. 1, pp. 1–15, Jan. 2021.

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2818–2826.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.* Amsterdam, The Netherlands: Springer, 2016, pp. 630–645.

[13] A. K. Gangwar and V. Ravi, "Diabetic retinopathy detection using transfer learning and deep learning," in *Evolution in Computational Intelligence*. Singapore: Springer, 2021, pp. 679–689.

[14] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification," *J. Ambient Intell. Hum. Comput.*, vol. 12, pp. 1–15, Jan. 2021.

[15] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, p. 3704, May 2021.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[17] O. Dekhil, A. Naglah, M. Shaban, M. Ghazal, F. Taher, and A. Elbaz, "Deep learning based method for computer aided diagnosis of diabetic retinopathy," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Dec. 2019, pp. 1–4.

[18] K. Shankar, Y. Zhang, Y. Liu, L. Wu, and C.-H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020.

[19] S. Majumder, Y. Elloumi, M. Akil, R. Kachouri, and N. Kehtarnavaz, "A deep learning-based smartphone app for real-time detection of five stages of diabetic retinopathy," *Proc. SPIE*, vol. 11401, Apr. 2020, Art. no. 1140106.

[20] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019.

[21] N. Sikder, M. Masud, A. K. Bairagi, A. S. M. Arif, A.-A. Nahid, and H. A. Alhumyani, "Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images," *Symmetry*, vol. 13, no. 4, p. 670, Apr. 2021.

[22] G. S., V. P. Gopi, and P. Palanisamy, "A lightweight CNN for diabetic retinopathy classification from fundus images," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102115.

[23] V. T, S. M, A. Kumaravel, and K. B, "Gabor filter and machine learning based diabetic retinopathy analysis and detection," *Microprocessors Microsyst.*, vol. 2020, Oct. 2020, Art. no. 103353.

[24] H.-Y. Tsao, P.-Y. Chan, and E. C.-Y. Su, "Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms," *BMC Bioinf.*, vol. 19, no. S9, pp. 111–121, Aug. 2018.

[25] S. Somasundaram and P. Alli, "A machine learning ensemble classifier for early prediction of diabetic retinopathy," *J. Med. Syst.*, vol. 41, no. 12, pp. 1–12, Dec. 2017.

[26] A. R. Rajanna, K. Aryafar, R. Ramchandran, C. Sisson, A. Shokoufandeh, and R. Ptucha, "Neural networks with manifold learning for diabetic retinopathy detection," 2016, *arXiv:1612.03961*.

[27] M. A. Al-Jarrah and H. Shatnawi, "Non-proliferative diabetic retinopathy symptoms detection and classification using neural network," *J. Med. Eng. Technol.*, vol. 41, no. 6, pp. 498–505, Aug. 2017.

[28] P. Costa, A. Galdran, A. Smailagic, and A. Campilho, "A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images," *IEEE Access*, vol. 6, pp. 18747–18758, 2018.

[29] A. M. Pour, H. Seyedarabi, S. H. A. Jahromi, and A. Javadzadeh, "Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization," *IEEE Access*, vol. 8, pp. 136668–136673, 2020.

[30] X. Zeng, H. Chen, Y. Luo, and W. Ye, "Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network," *IEEE Access*, vol. 7, pp. 30744–30753, 2019.

[31] G. Quellec, S. R. Russell, and M. D. Abramoff, "Optimal filter framework for automated, instantaneous detection of lesions in retinal images," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 523–533, Feb. 2011.

[32] Asia Pacific Tele-Ophthalmology Society (APTOS). (2019). *Messidor-Adcis*. Accessed: Sep. 11, 2021. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection

[33] (2010). *Messidor-Adcis*. Accessed: Oct. 18, 2021. [Online]. Available: https://www.adcis.net/en/third-party/messidor/

[34] C. P. Wilkinson, F. L. Ferris III, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, and J. T. Verdaguer, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, Sep. 2003.

[35] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 43, no. 1, pp. 1–8, Feb. 1997.

[36] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[37] S. Sarangi, M. Sahidullah, and G. Saha, "Optimization of data-driven filterbank for automatic speaker verification," *Digit. Signal Process.*, vol. 104, Sep. 2020, Art. no. 102795.

[38] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.

[39] S. Mittal, "A survey of FPGA-based accelerators for convolutional neural networks," *Neural Comput. Appl.*, vol. 32, no. 4, pp. 1109–1139, Feb. 2020.

[40] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3642–3649.

[41] G. Kovács, L. Tóth, D. Van Compernolle, and S. Ganapathy, "Increasing the robustness of CNN acoustic models using autoregressive moving average spectrogram features and channel dropout," *Pattern Recognit. Lett.*, vol. 100, pp. 44–50, Dec. 2017.

[42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[43] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, 2006.

[44] D. M. W. Powers, "What the F-measure doesn't measure: Features, flaws, fallacies and fixes," 2015, *arXiv:1503.06410*.
[45] L. V. R. Asuncion, J. X. P. De Mesa, P. K. H. Juan, N. T. Sayson, and A. R. D. Cruz, "Thigh motion-based gait analysis for human identification using inertial measurement units (IMUs)," in *Proc. IEEE 10th Int. Conf. Hum., Nanotechnol., Inf. Technol., Commun. Control, Environ. Manage. (HNICEM)*, Nov. 2018, pp. 1–6.
[46] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, *arXiv:2010.16061*.
[47] Y. Sasaki and R. Fellow, "The truth of the F-measure, Manchester: MIB-School of Computer Science," Univ. Manchester, Manchester, U.K., 2007, p. 25.
[48] D. J. Peres, C. Iuppa, L. Cavallaro, A. Cancelliere, and E. Foti, "Significant wave height record extension by neural networks and reanalysis wind data," *Ocean Model.*, vol. 94, pp. 128–140, Oct. 2015.
[49] V. Naralasetti, S. N. Shareef, and S. Hakak, "Blended multi-modal deep convnet features for diabetic retinopathy severity prediction," *Electronics*, vol. 9, no. 6, p. 914, 2020.
[50] C. I. Sánchez, M. Niemeijer, A. V. Dumitrescu, M. S. Suttorp-Schulten, M. D. Abràmoff, and B. van Ginneken, "Evaluation of a computer-aided diagnosis system for diabetic retinopathy screening on public data," *Investigative Ophthalmol. Vis. Sci.*, vol. 52, no. 7, pp. 4866–4871, Jun. 2011.
[51] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. M. P. Langlois, "Red lesion detection using dynamic shape features for diabetic retinopathy screening," *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1116–1126, Apr. 2015.
[52] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.
[53] H. Liu, K. Yue, S. Cheng, C. Pan, J. Sun, and W. Li, "Hybrid model structure for diabetic retinopathy classification," *J. Healthcare Eng.*, vol. 2020, pp. 1–9, Oct. 2020.
[54] V. Dondeti, J. D. Bodapati, S. N. Shareef, and N. Veeranjaneyulu, "Deep convolution features in non-linear embedding space for fundus image classification," *Rev. Intell. Artif.*, vol. 34, no. 3, pp. 307–313, 2020.
[55] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Diabetic retinopathy classification using a modified xception architecture," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2019, pp. 1–6.

MD. NAHIDUZZAMAN received the B.Sc. degree in computer science and engineering from the Rajshahi University of Engineering & Technology, Kazla, Rajshahi, Bangladesh, in 2018, where he is currently pursuing the M.Sc. degree. He is also serving as a Lecturer for the Department of Electrical and Computer Engineering, Rajshahi University of Engineering & Technology. He is interested in machine learning, deep learning, and pattern recognition.

MD. ROBIUL ISLAM received the B.Sc. degree in computer science and engineering from the Rajshahi University of Engineering & Technology, Kazla, Rajshahi, Bangladesh, in 2017, where he is currently pursuing the M.Sc. degree. He is also serving as a Lecturer for the Department of Electrical and Computer Engineering, Rajshahi University of Engineering & Technology. He is interested in segmentation, deep learning, and pattern recognition.

S. M. RIAZUL ISLAM (Member, IEEE) was with the University of Dhaka, Bangladesh, as an Assistant Professor, and a Lecturer with the Department of Electrical and Electronic Engineering, from 2005 to 2014. In 2014, he worked at Samsung Research and Development Institute Bangladesh, as a Chief Engineer with the Department of Solution Laboratory for Advanced Research. He worked at the Wireless Communications Research Center, Inha University, South Korea, as a Postdoctoral Fellow, from 2014 to 2017. From 2016 to 2017, he was also affiliated with Memorial University, Canada, as a Postdoctoral Fellow. He has been working as an Assistant Professor with the Department of Computer Science and Engineering, Sejong University, South Korea, since March 2017. His research interests include wireless communications, the Internet of Things, and applied artificial intelligence.

MD. OMAER FARUQ GONI received the B.Sc. degree in electrical and computer engineering from the Rajshahi University of Engineering & Technology, Kazla, Rajshahi, Bangladesh, in 2021. He is interested in deep learning, smart grid, and robotics.

MD. SHAMIM ANOWER (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of New South Wales, Australia, in 2012. He is currently serving as a Professor for the Department of Electrical and Electronic Engineering, Rajshahi University of Engineering & Technology. He has over 150 refereed publications, including journal articles (18 Q1 ranked in Scimago) and conference papers with H-index 20, I10-index 33, and 1205 citations. These cover the target sectors, namely energy and mining technology, quantum information, advanced digital, data science and ICT, and cyber security. He is a fellow of the Institute of Engineers Bangladesh (IEB).

KYUNG-SUP KWAK (Life Senior Member, IEEE) received the Ph.D. degree from the University of California. He was with Hughes Network Systems and the IBM Network Analysis Center, USA. He was with Inha University, South Korea, as a Professor. He was also the Dean of the Graduate School of Information Technology and Telecommunications and the Director of the UWB Wireless Communications Research Center. In 2008, he was an Inha Fellow Professor (IFP). He is currently an Inha Hanlim Fellow Professor and a Professor with the School of Information and Communication Engineering, Inha University. His research interests include UWB radio systems, wireless body area networks and u-health networks, and nano and molecular communications. In 2006, he was the President of the Korean Institute of Communication Sciences (KICS) and the Korea Institute of Intelligent Transport Systems (KITS), in 2009. He received the official commendations for achievements of UWB radio technology research and development from the Korean President, in 2009.