A Multi-Stage Detection of Diabetic Retinopathy in Fundus Images Using Convolutional Neural Network

Puneet Kumar[1], Salil Bharany[2], Ateeq Ur Rehman[3]*, Arjumand Bono Soomro[4], Mohammad Shuaib Mir[5], Yonis Gulzar[6]*

Department of CSE-Chandigarh Group of Colleges, Chandigarh Engineering College, Jhanjeri, Mohali, Punjab–140307, India[1]
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, India[2]
School of Computing, Gachon University, Seongnam-si 13120, Republic of Korea[3]
Department of Management Information Systems-College of Business Administration,
King Faisal University, Al-Ahsa 31982, Saudi Arabia[4, 5, 6]

***Abstract*—Diabetic Retinopathy (DRY) is a microvascular complication caused by diabetes mellitus, and it is one of the leading causes of blindness, especially in human adults. As the prevalence of this disease is growing exponentially, the screening of millions of people needs to be performed at a proliferating rate to diagnose the stage of the disease in its early stages. Highly advanced in the domain of technology, especially in artificial intelligence and its allied techniques, has come for the screening of DRY in photography to enhance the quality of life. This generates a bulk size of data that travels at high speed and cuts down on many human tasks. However, the techniques employed by the authors so far are quite expensive and time-consuming, and the prediction rate is insufficient to apply in a real-time scenario. This study offered a road for a deep learning-based fully automated system that helps to save manual disease diagnosis work and achieve disease detection in its very early stage using EfficientNetB3 (ENB3) Convolutional Neural Network (CNN) on DRY Fundus Images (FDI). In the suggested CNN, architectural variations and pre-processing techniques such as dimensionality reduction, global average pooling, and circular cropping are introduced alongside the Leaky ReLU (LR) activation function, Transfer Learning, and Reduce LROnPlateau technique, respectively. The accuracy of the proposed CNN classifier was 94.2% on training data, with a kappa score of 0.874, while it achieved a high level of accuracy at 96.7% on the testing data for DRY grading. Further, the evaluation results presented that the proposed model efficiently classifies the DRY stages for early disease detection.**

***Keywords*—Diabetic retinopathy; convolutional neural network; EfficientNetB3; fundus images; deep learning; transfer learning**

I. INTRODUCTION

Diabetes is growing exponentially and has been the primary reason for vision loss in human beings over the past few decades. Jan et al. [1] have pointed out that Diabetes Mellitus (DM) is present in a significant proportion of the global population, accounting for 382 million, which is expected to go even higher and reach new heights of somewhere around 592 million by 2025. It is further classified into two categories: Type-I and Type-II diabetes. The diseases are classified on the basis of the symptoms possessed by the human body, especially in the eyes. A significant impact on the eye is stated due to DM, known as Diabetic Retinopathy (DRY). Around 34.2% suffer from the DRY disease, among 382 million people who suffer from DM. In addition to it, 6.8% and 7% (approximately) have been detected with Diabetic Macular Edema (DME) and Proliferative Diabetic Retinopathy (PDRY), respectively [2]. These diseases are growing exponentially, especially in working-age adults and the aging populace. Therefore, the number of patients suffering from the DRY disease may be increased from 126.6 million to 191 million worldwide by 2030 [2]. The primary reason for people suffering from DRY is the working style of adults and not taking care of the preventive measures that control blood sugar. In general, detecting DRY in the early stage is a highly complex task, and people also feel asymptotic during the initial stage. However, in the later stages, many symptoms, like blurry vision, distortions, etc., are possessed with this disease. In the most severe stage, a human being may completely lose their eye-sight visualization. Therefore, the early prediction of the DRY disease is the most essential task to avoid the consequences in the later stage.

There are two broad classes of DRY: PDRY and Non-Proliferative Diabetic Retinopathy (NPDRY). Further, there are three eminent categories of NPDRY to detect the early stage of eye disease: Mild DRY, Moderate DRY, and Severe DRY [3]–[8]. The retinal fundoscopic illustration of different stages is depicted in Fig. 1 [9]. Afterwards, the classification diagram for these diseases is illustrated in Fig. 2 [3]. Moreover, the impact on human eyes during the distinguished DRY stages is presented in Fig. 3 [5].

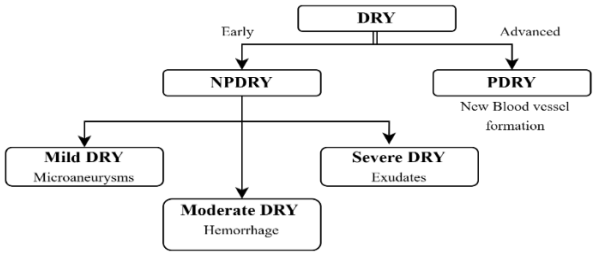

Fig. 1. DRY Classification.

A primary motivation to work in this field is the concern for human health [63], as it is evident that in the human body, one of the most important sensory organs is the eye. Over the last few years, eye diseases have grown exponentially in human beings [1]. Thus, it is necessary to establish a system for the

Corresponding Author: 202411144@gachon.ac.kr (A.U.R); ygulzar@kfu.edu.sa (Y.G.)

early diagnosis of the stage of such curable diseases to treat promptly. In the past decade, eye disease detection has also been extensively dependent on Deep Learning (DL) systems. Here, DL is said to identify critical patterns that may not be easily detectable by humans in visual data, making extensive analysis of complex visual data from DL [10]–[12][64]. One entails DRY, a disease that affects blood vessels within the retina, and blindness can easily result from it. DL models have been trained to detect signs of DRY by analyzing retinal images and identifying abnormalities in the blood vessels [13]–[15].

Age-related macular degeneration is another condition wherein the central part of the retina is damaged and can also cause blindness [16]–[18]. DL models are currently utilized for retinal images and processed for the detection of early signs, enabling earlier treatments and preventive measures for losing vision. Overall, using DL models in eye disease detection can potentially improve patient outcomes and significantly reduce the incidence of blindness. Now, the main research question is: "How can a multi-stage CNN architecture be effectively designed to optimize the detection accuracy and stage-wise classification of diabetic retinopathy in fundus images?"

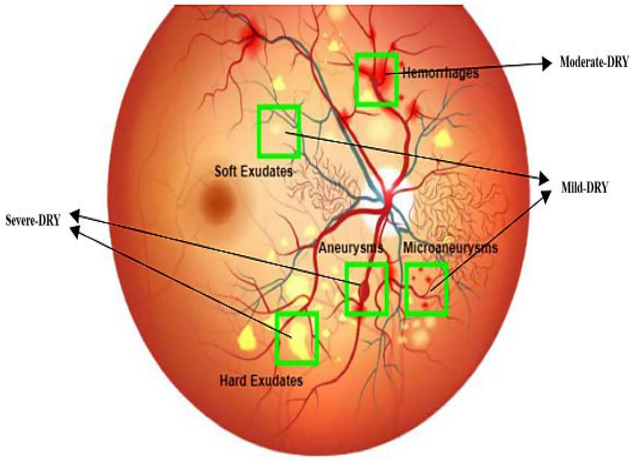

Fig. 2. Retinal fundoscopic illustration.

Therefore, the primary focus of this study is to classify the five stages of DRY disease in the most efficient manner, such that predicting early stages can be performed using the ocular eye disorders FDI to avert the possibility of blindness in human beings. Moreover, present research on this concern has shown that the CNN approaches are employed widely and ultimately outperform the various traditional and conventional approaches in distinguishing disease classification, segmentation, and prediction [19]–[24].

This study presents a novel multi-class detection system for DRY using an enhanced EfficientNetB3 CNN model. The key innovations include architectural modifications like dimensionality reduction with global average polling, circular cropping, and incorporation of LR activation function, Transfer Learning, and ReduceLROnPlateau technique. The model attained superior performance based on the training data of 94.2% accuracy and on a 0.874 kappa score, while the testing data resulted in 96.7% accuracy for the DRY grade-a system that outperformed earlier methods. The system is fully autonomous, eliminating the need for manual disease diagnosis, and excels at early-stage detection. Additionally, the model demonstrates strong potential for broader applications in biomedical image interpretation and medical decision-making while addressing previous limitations like expensive computation and insufficient prediction rates in real-time scenarios. The high accuracy and robust performance make it a considerable leap in automated DR detection and classification.

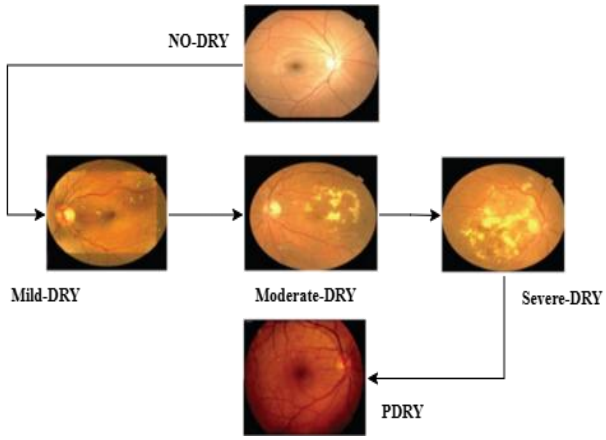

Fig. 3. Stages of DRY.

The organization of the rest of the study is as follows. In Section II, a background study of various state-of-the-art models, along with other relevant details related to various ML and DL approaches, is provided. Furthermore, this section contains an account of techniques and descriptions involved in grading DRY classification. The dataset description, preprocessing steps, methodology, and architectural diagram are described in the subsequent Sections III and IV. Section V brings together experimental settings, parameters for performance measurement, graphical results, and a detailed discussion. Finally, the future scope and conclusion of the research are represented in Section VI.

II. Literature Review

Eye diseases have been the primary cause of vision loss among human beings and have increased exponentially in the past few decades [1],[8]. Several research articles have been published to detect the early stage of eye disease to cure human life. This particular section is concerned with some of the approaches and techniques employed by the researcher in the prediction and classification of the grading of DRY in humans. These gradings are described as class-0 to class-4 as per the severity level, where class-0 represents the No DRY and class-4 represents the most severe DRY level.

Ryu et al. [28] put forward a model that is absolutely automatic and CNN-based for DRY detection and referable DRY from Optical Coherence Tomography Angiography (OCTA) scans. Furthermore, the validation of the model has been carried out using an external dataset, accompanied by four cross-validation techniques applied for model training. The proposed CNN model achieves classification accuracy between 91% to 95% for onset DRY and 91% to 98% for referable DRY. However, the dataset used for the work by the authors includes only 301 eye scan images, which was a major setback for the

proposed work. Math et al. [29] proposed a model for DRY detection based on segment-based learning.

Moreover, the CNN pre-trained model is stacked with an end-to-end segment-based learning method. The system was further validated using the Kaggle DRY detection dataset [30] and DIARET-DB1 [31] datasets, achieving 96.3% classification accuracy. A Synergic DL model has been proposed by Kathiresan et al. [32] presented a Synergic DL model to classify the DRY FDI. The first step of the authors was to do a noise-removal process, after which a histogram-based segmentation was applied to extract the relevant features from the images scanned for DRY. To validate the model, the authors used the MESSIDOR dataset [33]. Alyoubi et al. [34] reviewed 33 papers that classify eye disease using DL approaches. The author's main concern was to present the importance of the DL approaches to detect the DRY. Further, various data augmentation approaches were also described in the study to avoid overfitting.

Jiang et al. [35] suggested a technique that uses the Adaboost algorithm to classify the DRY using FDI. The authors used the demographically specific Chinese population dataset, which had 30,244 FDI. In the initial step, they resized all the images to (520, 520, 3) pixels, and then the augmentation steps were employed. The implementation part was divided into three subcategories: distribution, initialization, and combination of the illustrated model. An automated DRY classification system using four categories of a high-quality dataset of disease images is proposed by Zhang and his team [36]. The dataset contains 13,767 disease scans on which the data augmentation and histogram equalization techniques were initially employed. Moreover, the contrast stretch algorithm was incorporated to lighten the quality of dark images. Initially, authors finetuned Inception V3, Dense Net, ResNet-50, Inception ResNet V2, and Xception models. It was later that the authors made use of the R-Fully convolution Network and integration to discover the relation of the class labels with the classifier part for the study. Tymchenko et al. [37] presented a multitask learning approach using CNN to classify the DRY stages. Initially, to train the model, the Kaggle EyePACs dataset [30] was used, and later on, the MESSIDOR dataset [38] contained 1200 FDI, and the IDRiD Dataset [39] contained 413 scans merged into it. The further approaches took advantage of TL by accessing the weights of the pre-trained ImageNet model for the initialization of the encoder. The model was tested with the Kaggle APTOS 2019 dataset [40]. Papadopoulos et al. [41] suggested an ML-based approach to detect DRY FDI. The approach is purely based on multiple-instance learning. From the image patches, local information was extracted first and then clubbed with the attention mechanism. The Hough transformation, resizing, cropping, and various other techniques were employed in the pre-processing stage because this model achieved state-of-the-art accuracy. The concept of TL was initially employed by Gangwar et al. [42] in proposing a hybrid model of CNN Inception-ResNet-v2. The architecture was modified by adding the ensemble-based Inception block at the top of the structure. The suggested model outperforms the Google Net model results. Majumder et al. [43] presented a multitask model for DRY classification. This approach introduced a hybrid combination of classification and regression models. The extracted features achieved from the hybrid combination were further passed to the multi-layer perceptron network. For EyePACs and APTOS datasets, the desired model yielded kappa scores of 0.90 and 0.88, respectively. Alyoubi et al. [44] have proposed an ensemble model for the classification of DRY using the models CNN512 and YOLOv3. The model was trained and tested using the Kaggle APTOS 2019 dataset. The proposed model reached an accuracy of 89%.

As per the present literature review, various ML and DL techniques have been investigated by the researchers to detect and classify DRY from retinal images [45]–[49]. However, despite the significant progress achieved in this area, several challenges and limitations still exist. One of them is the reduced availability of labeled data, which limits training DL models. Additionally, many models used in disease classification only had a few layers modified with the ReLU activation function. This activation function sets any negative input values to zero, making some neurons inactive due to a zero slope. ReLU may explode during training, affecting the model's convergence and sometimes leading to poor performance and less accurate results.

Moreover, the existing approaches often lack interpretability, making it challenging to understand the underlying decision-making process of these models. Furthermore, the performance of the existing models on diverse datasets remains inconsistent. Thus, there is an urgent need to develop highly robust and interpretable models that can also bridge the gaps in detection and classification as far as DRY is concerned. The proposed model outperforms previous models in several key ways. Firstly, it utilizes a more complex neural network architecture, incorporating additional layers and techniques like dimensionality reduction, global average pooling, circular cropping, etc., to better capture and learn from the underlying patterns in the data. This results in improved accuracy, kappa score, and predictive power compared to the various existing models. Additionally, the model leverages advanced optimization techniques such as Adam optimizer with LeakyRelu activation function to improve convergence and avoid getting stuck in local minima during training. Secondly, the proposed model incorporates more diverse and representative training data, allowing it to better generalize to unseen examples and minimize overfitting. Achieving this entails using augmentation and TL techniques that increase the diversity and quantity of training data available. The improvements make the current proposed DL model a more powerful and reliable tool to tackle such real-time complex problems.

III. Materials and Methods

This section presents an overview of the dataset employed in this research work, as well as descriptive pre-processing and augmentation techniques.

A. Dataset Description

The largest publicly available FDI of DRY has been taken from the Kaggle repository to perform the experimental analysis [30]. The dataset is actually made up of 88,702 images, a total of which 35,126 images have been labeled, whereas the remaining 53,576 images have not been labeled. This dataset was selected due to its popularity in DRY studies. Also, it

contains extremely well-labeled fundus images with different levels of disease severity. Further, the dataset is divided into five classes based on eye disease severity level, namely: No DRY, Mild DRY, Moderate Dry, Severe Dry, and Proliferative DRY, which are represented as Class-0, 1, 2, 3, and Class-4, respectively. The class-wise instances of eye diseases are presented in Fig. 4.

The only labeled images are used for eye disease diagnosis procedures in the study aimed entirely at classifying various phases of DRY. A detailed account of different classes in the dataset is presented in Table I and Fig. 5.

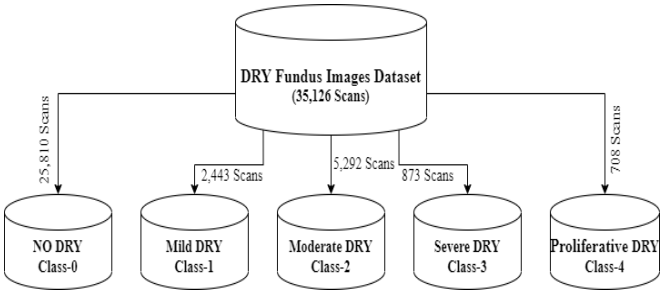

Fig. 4. Class-wise instances of eye disease.

TABLE I. DATASET DESCRIPTION- DIFFERENT CLASS INSTANCES

Categories	No-DRY Normal (Class-0)	Mild DRY (Class-1)	Moderate DRY (Class-2)	Severe DRY (Class-3)	Proliferative DRY (Class-4)	Total Scan
Total Images	25810	2443	5292	873	708	35126
Percentage %	73.48%	6.95%	15.07%	2.49%	2.02%	100%
Training Set	16536	1563	3359	558	453	22469
Validation Set	5155	489	1088	175	142	7049
Test Set	4119	391	845	140	113	5608

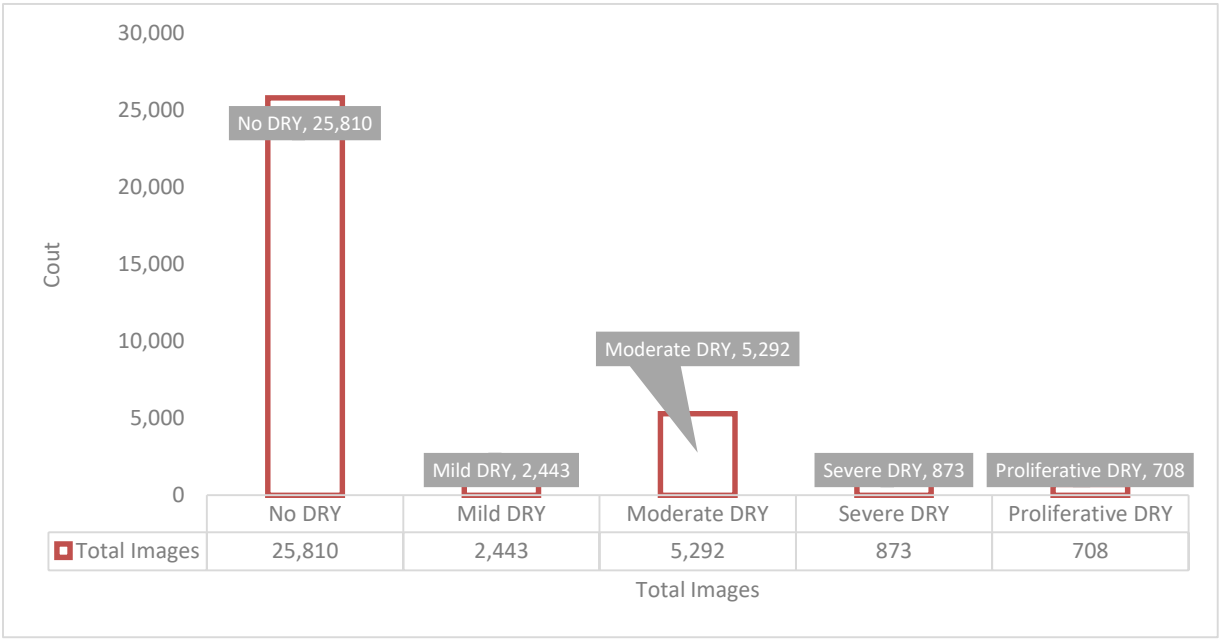

Fig. 5. Description of a collected dataset for eye disease.

B. Data Pre-processing

After collecting the dataset, a pre-processing step was performed. As the size of the original image is very large (say, 3000 x 2000 pixels on average), therefore in the initial phase, all the images are resized to (320, 320, 3), i.e., (width, height, channel), before passing to the model. The size of the images is capable of avoiding feature loss and is suitable for efficient training of the proposed model. The illustration of the images after resizing is presented in Fig. 6. Afterward, the region of interest is identified, and circular cropping is applied to the dataset. It helps to remove the background area and enables the model to work on the specified region of interest without any significant information loss.

C. Data Augmentation

Data augmentation is also applied using Keras's image-data generator, which includes a zoom range of 0.3, a brightness range of 0.5, a rotation range of 360°, etc. The results achieved after applying the data augmentation are presented in Fig. 7. Moreover, to resolve the overfitting issue, the model loss was optimized using a learning rate of 1e-4 and decay of 1e-6.

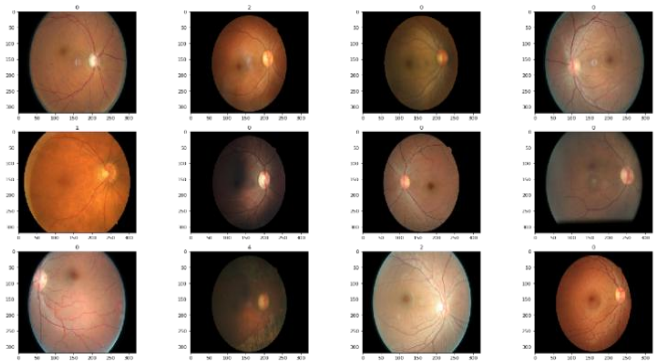

Fig. 6. Images after pre-processing of eye diseases.

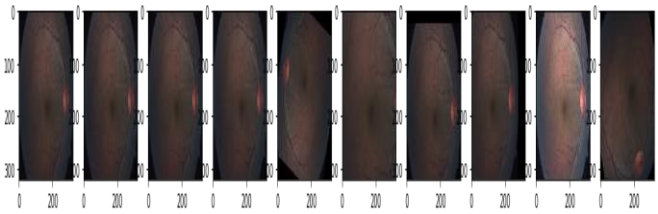

Fig. 7. Augmented images of eye diseases (after circular cropping).

D. Methods

CNN belongs to a specific class of neural networks that have been created to excel at recognizing and categorizing images, leading to impressive outcomes in this domain. Their ability to autonomously grasp intricate and endure attributes directly from the raw images sets CNN apart from conventional methods, eliminating the need for laborious manual feature extraction [50]–[53]. When applied to identifying different stages of DRY, CNN has demonstrated its superiority by delivering better results compared to traditional feature extraction techniques [54]–[59]. A standard CNN configuration primarily comprises: a) convolution layers, b) pooling layers, and c) fully connected layers. On the other hand, the ENB3 model is reported to be a strong and significantly advanced neural network model. It has a good compromise between computation and performance to make it an efficient model of choice for eye disease classification. ENB3 is a CNN architecture that reduces the size of the parameters by using the compound coefficient. Therefore, the depth, width, and resolution dimensions are uniformly scaled down, further increasing this state-of-the-art model's efficiency and accuracy.

The advantages of employing this CNN architecture are prominently leveraged within the proposed model. This architecture supplements the ENB3 model with additional neural network layers, which include global average pooling, dropout, and dense layers. These layers are augmented with the LR activation function, effectively addressing the challenge of inactive neurons that can arise during model training. In addition, hyperparameter optimization is carried out using the ReduceLROnPlateau technique. This improvement considerably boosts the proposed model from the aspect of accuracy and efficiency while reducing validation loss in each epoch of model training.

IV. Proposed Methodology

The successful execution of research endeavors rests upon a well-defined methodology. Therefore, the proposed EfficientNetB3_Model methodology work and architectural formulation are discussed in this section. By stacking the ENB3 model with multiple layers of deep neural networks using the LR activation function and the ReduceLROnPlateau technique, the model automatically classifies and grades the DRY's retinal FDI.

A. System Model

ENB3 is a CNN architecture that reduces the size of the parameters by using the compound coefficient. In this section, a step-by-step explanation of the model approach is provided.

Step 1: Resize all Images to (320, 320, 3), i.e., (width, height, channel)

Step 2: Identify the region of interest and create a circular crop around the image center, where all circles center on the x-axis.

$$Y^2.Y'^2 + Y^2 = R^2 \quad (1)$$

In the given context, Y denotes the y-axis and Y' represents the differentiation, and R represents the radius.

Step 3: Data Augmentation Phase:

The Keras Image data generator is utilized to apply data augmentation.

- ZoomRange - 0.3
- HorizontalFlip - True
- VerticalFlip - True
- BrightnessRange - (0.5, 2)
- RotationRange - 360º
- ZoomRange - (0.65, 1)

This phase significantly increases the data diversity.

Step 4: The ENB3 model was called, and the LR activation function was applied. The pre-processed images were passed as input for feature extraction.

Step 5: Fetch output from the final block of ENB3 and apply the two-dimensional Global Average Pooling (f_{Gavg}):

$$f_{Gavg}(x) = \frac{1}{N}\sum_{i=1}^{n} x_i \quad (2)$$

where, x is the vector consisting of activation values from a rectangular area of N pixels.

Step 6: A Dropout with a rate of 0.5 was added to the model.

$$y = activation\ (dot\ (input, kernel) + b)\ ^\circ m \quad (3)$$

where, output is represented by y, the activation function is represented by the keyword activation, inputs and weights of product represented by the dot, input data is represented by the input keyword, weights and bias are represented by kernel and bias represent the element-wise multiplication, m is the $Bernouli\ (p)$, and p is the dropout probability.

Step 7: Add BatchNormalization Layer

Step 8: A dense layer using the Sigmoid activation function is integrated for the interface.

$$z = activation\,(dot\,(input, kernel) + b) \quad (4)$$

$$Sigmoid(z_i) = \frac{1}{(1+exp^{-z})} \quad (5)$$

where, z_i represents the input vector.

Step 9: The model was compiled using the Adam optimizer, with a learning rate of 1e-4 and decay of 1e-6.

$$w_t = w_{t-1} - \delta \frac{M_t}{\sqrt{V_t}+\epsilon} \quad (6)$$

where, w is the model weight, δ is the step size, β_1 *and* β_2 are the hyperparameters, t is the time stamp, m is the mean, and v is the variance represented as:

$$M_t = {}^{m_t}/_{1-\beta_1^t} \quad (7)$$

$$V_t = {}^{v_t}/_{1-\beta_2^t} \quad (8)$$

Step 10: Apply the binary cross-entropy loss function L_{BCE}

$$L_{BCE} = -\frac{1}{out.size} \sum_{i=1}^{out.size} \hat{z_i} * \log z_i + (1 - z_i) * \log(1 - \hat{z_i}) \quad (9)$$

In the given context, $\hat{z_i}$ is the indicator of the ith scaler value in the model output, whereas z_i is the indicator of the appropriate corresponding target value, and out.size is the number that indicates the model output scalar values.

Step 11: Fine-tune the model by using ReduceLROnPlateau.

$$LR_i = init_{LR} * (\frac{mxt_{LR}}{init_{LR}})^{i/n} \quad (10)$$

where, $init_{LR}$ is-initial learning rate, mxt_{LR} -is the maximum learning rate, n-is the no. of iterations, and i-stands for i_{th} min-batch.

B. Architecture and Working

The proposed ENB3 model is formulated by stacking the EfficientNetB3 model and different deep neural network layers. A detailed illustration of the architecture is described in Fig. 8. In this, the ENB3 model weights are accessed using the TL concept. Therefore, rather than training a model from scratch, using pre-trained model weights highly increases the speed at which the model operates and minimizes the model's overall training time [60], [61].

Another thing that is used to stack the network structure is to add layers, which include global average pooling and dropout, batch normalization, and finally dense layers. These layers are incorporated with the LR activation function, which evades the problems of overfitting that generally occur during the model's training and helps achieve model generalization. Afterward, the model finetuning is performed using the ReduceLROnPlateau, which further improves the model validation loss during each epoch of model training.

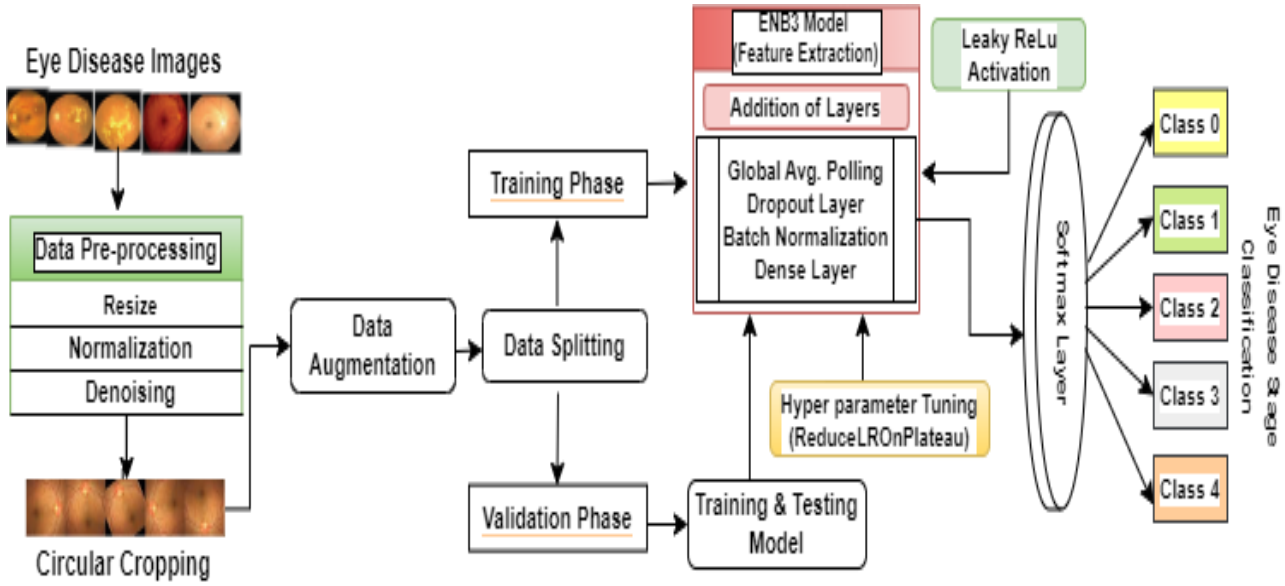

Fig. 8. Proposed ENB3_model architecture.

V. Experimental Result Analysis

This research phase involved hands-on experimentation to classify and grade the DRY's retinal FDI automatically. Therefore, in this section, the experimental setup, performance metrics utilized to analyze the model authentication, and analysis of results are discussed.

A. Experimental Setup

The proposed model is run within the infrastructure of a 64-bit Windows operating system version 10 and with an installed memory of 16GB. A GPU has 2560 CUDA cores (NVIDIA Tesla T4) forms part of this computational setup. Moreover, Keras, a package of deep learning with TensorFlow at the backend, is used for the implementation work [62].

B. Performance Metrics

The performance of the suggested model is evaluated according to a set of accuracy, specificity, precision, recall, kappa value, and other performance metrics. These metrics are crucial for assessing the model's effectiveness in accurately classifying data and identifying potential areas for improvement. Eq. (11) to Eq. (15) represent these performance metrics, quantitatively measuring the model's performance.

$$Precision = \frac{TP}{(TP+FP)} \quad (11)$$

$$Recall\ (or\ Senstivity) = \frac{TP}{(TP+FP)} \quad (12)$$

$$Accuracy = \frac{(TP+TN)}{(TP+FN+TN+FP)} \quad (13)$$

$$F1\ Score = 2 * \frac{(Precision*Recall)}{(Precision+Recall)} \quad (14)$$

$$Specificity = \frac{(TN)}{(TN+TP)} \quad (15)$$

- TruePositive (TP) indicates that the instances are correctly classified to DRY stages, which means the results attained from the model and the results available in the training set are the same.
- TrueNegative (TN) indicates that the instances are correctly classified as No DRY, which means the patient has no disease, and the model result is also negative.
- FalsePositive (FP) or Type-1Error indicates that the patient has no DRY, but the result predicted by the model is positive.
- FalseNegative (FN) or Type-2Error indicates that the patient has DRY, but the result predicted by the model is negative.

C. Results and Discussion

This section compares different DRY disease stages predicted by the proposed model with normal eye scans. Furthermore, the results generated by the proposed model have been explained using the various performance metrics. Later, the outcome generated by the model is compared with other state-of-the-art models. The model performed excellently; however, future scans of DRY are recommended for validating the performance further, which can allow the proposed model to be used in a real-time situation. ENB3 is a CNN architecture that reduces the size of the parameters by using the compound coefficient. Therefore, the depth, width, and resolution dimensions are uniformly scaled down, further increasing this state-of-the-art model's efficiency and accuracy. The benefits of this CNN architecture are further employed significantly in the proposed model. In this architecture, some additional neural network layers, such as global average pooling, dropout, and dense layers, are stacked with the ENB3 model. Further, these layers are incorporated with the LR activation function to resolve dead neurons during the model training. Afterward, ReduceLROnPlateau performs the hyperparameter tuning. This improves the proposed model's accuracy, efficiency, and validation loss during each epoch of the model training. The model's training is done only for forty epochs, as no further improvement in the loss was observed later on. The results achieved by the proposed model are illustrated graphically in Fig.s 9, 10, and 11. The training accuracy and validation accuracy of the suggested model are depicted in Fig. 9. The learning path of the proposed model is explained through a curves graphically, and illustrates that the model has learned quite well and no overfitting occurred. The accuracy attained by the model is 94.2% during the training phase, which is outstanding compared to the results attained by different models in past studies. All performance measures and parameters relating to the ENB3 model have been described uniformly in Table II.

TABLE II. PARAMETRIC PERFORMANCE OF ENB3 MODEL

S. No	Performance Measure and Parameters	
1	Model	EfficientNetB3
2	Epochs	40
3	Accuracy	0.942
4	Sensitivity	0.958
5	Specificity	0.90
6	Precision	0.959
7	F1 Score	0.958
8	Kappa Value	0.874
9	Training Loss	0.149
10	Validation Accuracy	0.93
11	Validation Loss	0.176
12	Learning Rate	1.00E-04
13	Batch Size	32

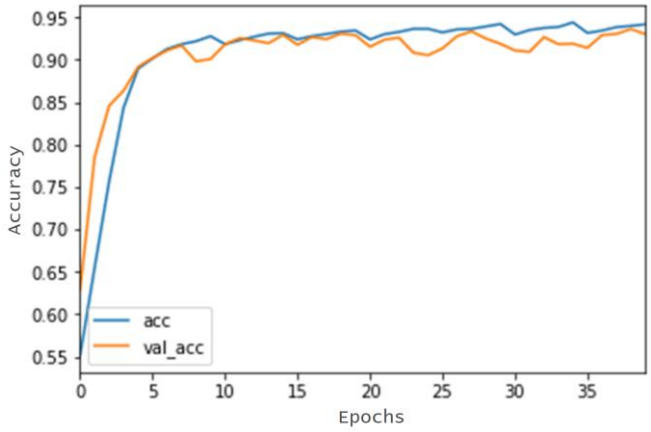

Fig. 9. ENB3 Model training accuracy.

The training and validation losses of the ENB3 model are illustrated in Fig. 10. Moreover, the learning capability of the model is also shown through curves in this figure. In the figure shown, the training loss drops exponentially during the first ten epochs, after which the decline in loss becomes gradual. Similarly, there is an exponential drop in validation loss in the first twelve epochs, but the subsequent decrease in loss is gradual. The model loss came out to be 0.149 and 0.176 during training and validation, respectively. Hence, it concludes that the model learns quite efficiently.

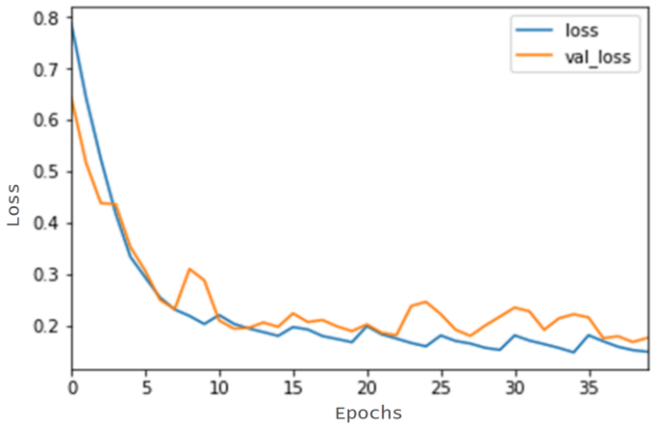

Fig. 10. ENB3 Model training loss.

D. Analysis

Another significant criterion of the kappa score deals with understanding the behavior of the model. Owing to the highly imbalanced nature of data, it provides a more realistic picture of the model's performance. In Fig. 11, the kappa value score obtained at each epoch is plotted against the different epochs during the training of the model. After 40 epochs in training mode, the model achieved a kappa score of 0.874.

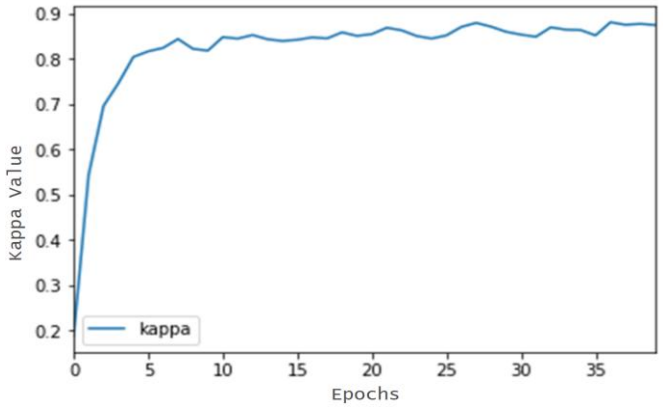

Fig. 11. ENB3 Model kappa score.

Furthermore, testing the model is another significant component to bring forth the performance of the suggested model in a real-time situation. The testing set's scans are then passed to the model in order to test the model. As the dataset is highly unbalanced and four different categories of DRY disease stages are present in the set, the class-wise performance evaluation results are described using the confusion matrix, as illustrated in Fig. 12.

The result of the confusion matrix clearly explained that the classification of all disease stages is done very efficiently by the ENB3 model. However, the model is slightly confused in predicting the DRY stages 2 and 3. The model tested on this test set has shown an accuracy of 96.67 % with a loss figure of 0.899. An explanation of the testing results is presented in Table III and IV.

Table V reveals the performance of the proposed model compared with existing models on DRY's detection [25]–[27].

A new method was introduced by Lin et al. [25] to improve DRY detection from retinal photographs by first converting them into entropy images. The study by Qummar et al. [4] presented a DL ensemble approach for DRY detection, which achieved good results. Shanthi et al. [26] proposed a modified AlexNet architecture for the classification of DRY images that outperformed previous methods. Another work takes up a modified AlexNet architecture with improved DRY image classification over previous studies. A classification system has been developed by Samah et al. [27] which diagnoses DRY using an enhanced image and CNN. The proposed study achieved remarkable results that portray the efficiency of the method proposed. Afterwards, the study by Liu et al. [22] which created a novel DL approach for DRY detection using a symmetric CNN. However, the result achieved from the proposed model shows that the ENB3 model surpasses the existing ensemble, non-ensemble, and state-of-the-art approaches. Broadly speaking, the work indicates the progress of the ENB3 method for detecting DRY stages accurately and efficiently. The results drawn by this study will provide insight into the formulation of reliable, effective diagnostic systems pertaining to DRY detection using the ENB3 model. Future research directions should continue to emerge through the exploration of upgrades to the methods for improved detection and diagnosis of DRY.

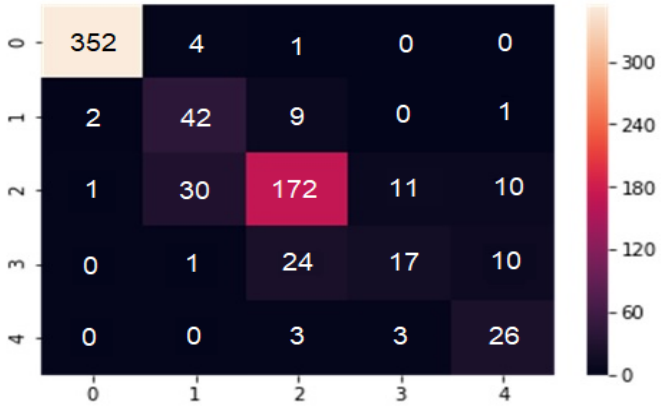

Fig. 12. Confusion matrix for eye disease stage classification.

TABLE III. PARAMETRIC PERFORMANCE OF ENB3 MODEL ON TESTING DATA

Label	Precision	F1-Score	Recall	Support
0	0.99	0.99	0.99	357
1	0.55	0.64	0.78	54
2	0.82	0.79	0.77	224
3	0.55	0.4	0.31	54
4	0.53	0.64	0.81	32

TABLE IV. TESTING RESULTS OF EYE DISEASE CLASSIFICATION

Model	Accuracy	Loss	Sensitivity	Specificity	Precision	F1 Score
EfficientNetB3	0.967	0.089	0.972	0.944	0.988	0.979

TABLE V. COMPARISON OF RESULTS (MULTI-STAGE CLASSIFICATION OF DRY)

Authors	Proposed Model	Accuracy	Sensitivity	Specificity
Lin et al. (2018) [25]	CNN	85.10%	80.2%	87.2%
Qummar et al. (2019) [4]	DL Ensemble Models	80.80%	51.50%	86.72%
Shanthi et al. (2019) [26]	CNN & Modified Alexnet	93.1%	93.2%	93%
Samah et al. (2019) [27]	CNN	92.80%	Not Available	Not Available
Liu et al. (2021) [22]	Deep Symmetric CNN	93.60%	93.70%	92.50%
ENB3 (Proposed Model)		**94.20%**	**95.80%**	**90.00%**

VI. CONCLUSION AND FUTURE ASPECTS

A deep learning-based eye disease DRY's detection approach using the ENB3 model is presented in this study. The proposed model automatically classifies and grades the DRY's retinal FDI. The major setback of the various existing models is their lack of accuracy. In this study, some architectural changes have been employed in the existing ENB3 CNN model that enhanced the efficiency and enriched the accuracy of the DRY's stages classification by using the FDI. Therefore, initially using the concept of TL, the weights of the pre-trained ENB3 model are accessed, and later on, pre-processing techniques like dimensionality reduction with global average pooling and circular cropping have been incorporated with the LR activation function. Lastly, the model is fine-tuned by employing the ReduceLROnPlateau technique, which extracts significant high-level features. Moreover, this concept helps to reduce complexity, avoid overfitting in the model, and extract discriminant features from the DRY FDI. In addition, disease stages are predicted on an imbalanced Kaggle dataset to validate the proposed model's performance. The results achieved from the proposed model indicate that the model surpasses all existing ones, whether ensemble or non-ensemble, and includes state-of-the-art methods. The proposed model achieved up to 94.2 % and 96.7% classification accuracy while grading DRY's stages in the training and testing sets, respectively. Although the proposed work is an important milestone in this area, there are some limitations, like the potential for the model to suffer from interpretability issues. DL models are generally regarded as black boxes, making it challenging to understand how predictions actually form. Moreover, alternative approaches or models may achieve comparable or superior performance to the proposed method, which should be explored in future research.

FUNDING STATEMENT

This work was supported by the Deanship of Scientific Research, the Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia under the project KFU251557.

Data Availability Statement: The data employed in the design and production cum layout of a manuscript is Available: https://www.kaggle.com/competitions/diabetic-retinopathy-detection/overview

Conflicts of Interest: The authors declare that they have no conflicts of interest to disclose in relation to this study.

REFERENCES

[1] S. Jan, I. Ahmad, S. Karim, Z. Hussain, M. Rehman, and M. A. Shah, "Status of diabetic retinopathy and its presentation patterns in diabetics at ophthalomogy clinics," J. Postgrad. Med. Inst., vol. 32, no. 1, pp. 24–27, 2018.

[2] Y. Zheng, M. He, and N. Congdon, "The worldwide epidemic of diabetic retinopathy," Indian J. Ophthalmol., vol. 60, no. 5, p. 428, 2012, doi: 10.4103/0301-4738.100542.

[3] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey," IEEE Access, vol. 10, pp. 28642–28655, 2022, doi: 10.1109/ACCESS.2022.3157632.

[4] S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," IEEE Access, vol. 7, pp. 150530–150539, 2019, doi: 10.1109/ACCESS.2019.2947484.

[5] J. De Fauw, "Detecting diabetic retinopathy in eye images," 2015. https://defauw.ai/diabetic-retinopathy-detection/ (accessed Jan. 21, 2022).

[6] L. Math and R. Fatima, "Adaptive machine learning classification for diabetic retinopathy," Multimed. Tools Appl., vol. 80, no. 4, pp. 5173–5186, Feb. 2021, doi: 10.1007/s11042-020-09793-7.

[7] N. George, L. Shine, A. N, B. Abraham, and S. Ramachandran, "A two-stage CNN model for the classification and severity analysis of retinal and choroidal diseases in OCT images," Int. J. Intell. Networks, vol. 5, pp. 10–18, 2024, doi: 10.1016/j.ijin.2024.01.002.

[8] E. M. F. El Houby, "Using transfer learning for diabetic retinopathy stage classification," Appl. Comput. Informatics, Oct. 2021, doi: 10.1108/ACI-07-2021-0191.

[9] P. Kumar, R. Kumar, and M. Gupta, "Deep Learning Based Analysis of Ophthalmology: A Systematic Review," EAI Endorsed Trans. Pervasive Heal. Technol., vol. 7, no. 29, 2021, doi: 10.4108/eai.10-9-2021.170950.

[10] A. Bali and V. Mansotra, "Analysis of Deep Learning Techniques for Prediction of Eye Diseases: A Systematic Review," Arch. Comput. Methods Eng., vol. 31, no. 1, pp. 487–520, Jan. 2024, doi: 10.1007/s11831-023-09989-8.

[11] X. Luo, J. Li, M. Chen, X. Yang, and X. Li, "Ophthalmic Disease Detection via Deep Learning With a Novel Mixture Loss Function," IEEE J. Biomed. Heal. Informatics, vol. 25, no. 9, pp. 3332–3339, Sep. 2021, doi: 10.1109/JBHI.2021.3083605.

[12] M. Buric, S. Grozdanic, and M. Ivasic-Kos, "Diagnosis of ophthalmologic diseases in canines based on images using neural networks for image segmentation," Heliyon, vol. 10, no. 19, p. e38287, Oct. 2024, doi: 10.1016/j.heliyon.2024.e38287.

[13] D. Maji and A. A. Sekh, "Automatic Grading of Retinal Blood Vessel in Deep Retinal Image Diagnosis," J. Med. Syst., vol. 44, no. 10, p. 180, Oct. 2020, doi: 10.1007/s10916-020-01635-1.

[14] K. Balasubramanian and N. P. Ananthamoorthy, "RETRACTED ARTICLE: Robust retinal blood vessel segmentation using convolutional neural network and support vector machine," J. Ambient Intell. Humaniz. Comput., vol. 12, no. 3, pp. 3559–3569, Mar. 2021, doi: 10.1007/s12652-019-01559-w.

[15] A. Krestanova, J. Kubicek, and M. Penhaker, "Recent Techniques and Trends for Retinal Blood Vessel Extraction and Tortuosity Evaluation: A Comprehensive Review," IEEE Access, vol. 8, pp. 197787–197816, 2020, doi: 10.1109/ACCESS.2020.3033027.

[16] T. J. Heesterbeek, L. Lorés-Motta, C. B. Hoyng, Y. T. E. Lechanteur, and A. I. den Hollander, "Risk factors for progression of age-related macular degeneration," Ophthalmic Physiol. Opt., vol. 40, no. 2, pp. 140–170, Mar. 2020, doi: 10.1111/opo.12675.

[17] M. J. Ammar, J. Hsu, A. Chiang, A. C. Ho, and C. D. Regillo, "Age-related macular degeneration therapy: a review," Curr. Opin. Ophthalmol., vol. 31, no. 3, pp. 215–221, May 2020, doi: 10.1097/ICU.0000000000000657.

[18] C. J. Thomas, R. G. Mirza, and M. K. Gill, "Age-Related Macular Degeneration," Med. Clin. North Am., vol. 105, no. 3, pp. 473–491, May 2021, doi: 10.1016/j.mcna.2021.01.003.

[19] Puneet, R. Kumar, and M. Gupta, "Optical coherence tomography image based eye disease detection using deep convolutional neural network," Heal. Inf. Sci. Syst., vol. 10, no. 1, p. 13, Dec. 2022, doi: 10.1007/s13755-022-00182-y.

[20] Olorunfemi, B.O., Ogunde, A.O., Almogren, A. et al. Efficient diagnosis of diabetes mellitus using an improved ensemble method. Sci Rep 15, 3235 (2025). https://doi.org/10.1038/s41598-025-87767-1

[21] Shandilya, G., Gupta, S., Almogren, A. et al. Enhancing advanced cervical cell categorization with cluster-based intelligent systems by a novel integrated CNN approach with skip mechanisms and GAN-based augmentation. Sci Rep 14, 29040 (2024). https://doi.org/10.1038/s41598-024-80260-1.

[22] Rani, S., Memoria, M., Almogren, A. et al. Deep learning to combat knee osteoarthritis and severity assessment by using CNN-based classification. BMC Musculoskelet Disord 25, 817 (2024). https://doi.org/10.1186/s12891-024-07942-9.

[23] 41. Maryam Shabbir, Zobia Suhail, Nida Hafeez, Najmus Saqib, Muhammad Farooq, Sghaier Guizani, Ateeq Ur Rehman*, Habib Hamam "Prostate Segmentation in MRI Images using Transfer Learning based Mask R-CNN" in current medical imaging, 2024, 20, e15734056305021. DOI: 10.2174/0115734056305021240603114137.

[24] D. Kumar and V. Kukreja, "Deep learning in wheat diseases classification: A systematic review," Multimed. Tools Appl., vol. 81, no. 7, pp. 10143–10187, Mar. 2022, doi: 10.1007/s11042-022-12160-3.

[25] G.-M. Lin et al., "Transforming Retinal Photographs to Entropy Images in Deep Learning to Improve Automated Detection for Diabetic Retinopathy," J. Ophthalmol., vol. 2018, pp. 1–6, Sep. 2018, doi: 10.1155/2018/2159702.

[26] T. Shanthi and R. S. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," Comput. Electr. Eng., vol. 76, pp. 56–64, Jun. 2019, doi: 10.1016/j.compeleceng.2019.03.004.

[27] A. H. Abu Samah, F. Ahmad, M. K. Osman, M. Idris, N. M. Tahir, and N. A. Abd. Aziz, "Classification of Pathological Signs for Diabetic Retinopathy Diagnosis using Image Enhancement Technique and Convolution Neural Network," in 2019 9th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Nov. 2019, pp. 221–225, doi: 10.1109/ICCSCE47578.2019.9068538.

[28] G. Ryu, K. Lee, D. Park, S. H. Park, and M. Sagong, "A deep learning model for identifying diabetic retinopathy using optical coherence tomography angiography," Sci. Rep., vol. 11, no. 1, pp. 1–9, 2021, doi: 10.1038/s41598-021-02479-6.

[29] L. Math and R. Fatima, "Adaptive machine learning classification for diabetic retinopathy," Multimed. Tools Appl., vol. 80, no. 4, pp. 5173–5186, 2021, doi: 10.1007/s11042-020-09793-7.

[30] C. H. Foundation., "Diabetic Retinopathy Detection 'Identify signs of diabetic retinopathy in eye images,'" 2015. Accessed: Jan. 12, 2022. [Online]. Available: https://www.kaggle.com/competitions/diabetic-retinopathy-detection/overview.

[31] T. Kauppi et al., "the DIARETDB1 diabetic retinopathy database and evaluation protocol," in Procedings of the British Machine Vision Conference 2007, 2007, vol. 10, no. 02, pp. 15.1-15.10, doi: 10.5244/C.21.15.

[32] K. Shankar, A. R. W. Sait, D. Gupta, S. K. Lakshmanaprabu, A. Khanna, and H. M. Pandey, "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," Pattern Recognit. Lett., vol. 133, pp. 210–216, 2020, doi: 10.1016/j.patrec.2020.02.026.

[33] M. Consortium, "Computer Vision experts in Image Acquisition / Processing / Artificial Intelligence." https://www.adcis.net/en/third-party/messidor/ (accessed Mar. 13, 2022).

[34] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," Informatics Med. Unlocked, vol. 20, p. 100377, 2020, doi: 10.1016/j.imu.2020.100377.

[35] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, and W. Qian, "An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2019, pp. 2045–2048, doi: 10.1109/EMBC.2019.8857160.

[36] W. Zhang et al., "Automated identification and grading system of diabetic retinopathy using deep neural networks," Knowledge-Based Syst., vol. 175, pp. 12–25, Jul. 2019, doi: 10.1016/j.knosys.2019.03.016.

[37] B. Tymchenko, P. Marchenko, and D. Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection," Mar. 2020, [Online]. Available: https://arxiv.org/abs/2003.02261.

[38] L. Giancardo et al., "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," Med. Image Anal., vol. 16, no. 1, pp. 216–226, Jan. 2012, doi: 10.1016/j.media.2011.07.004.

[39] P. Porwal et al., "Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research," Data, vol. 3, no. 3, p. 25, Jul. 2018, doi: 10.3390/data3030025.

[40] A. P. T.-O. S. (APTOS), "APTOS 2019 Blindness Detection." https://www.kaggle.com/c/aptos2019-blindness-detection (accessed Mar. 11, 2022).

[41] A. Papadopoulos, F. Topouzis, and A. Delopoulos, "An interpretable multiple-instance approach for the detection of referable diabetic retinopathy in fundus images," Sci. Rep., vol. 11, no. 1, p. 14326, Jul. 2021, doi: 10.1038/s41598-021-93632-8.

[42] A. K. Gangwar and V. Ravi, "Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning," in Advances in Intelligent Systems and Computing, vol. 1176, Springer Singapore, 2021, pp. 679–689.

[43] S. Majumder and N. Kehtarnavaz, "Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy," IEEE Access, vol. 9, pp. 123220–123230, 2021, doi: 10.1109/ACCESS.2021.3109240.

[44] 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, pp. 1–22, 2021, doi: 10.3390/s21113704.

[45] L. Abdel-Hamid, "Retinal image quality assessment using transfer learning: Spatial images vs. wavelet detail subbands," Ain Shams Eng. J., vol. 12, no. 3, pp. 2799–2807, Sep. 2021, doi: 10.1016/j.asej.2021.02.010.

[46] A. Ali et al., "Machine Learning Based Automated Segmentation and Hybrid Feature Analysis for Diabetic Retinopathy Classification Using Fundus Image," Entropy, vol. 22, no. 5, p. 567, May 2020, doi: 10.3390/e22050567.

[47] A. Raj, A. K. Tiwari, and M. G. Martini, "Fundus image quality assessment: survey, challenges, and future scope," IET Image Process., vol. 13, no. 8, pp. 1211–1224, Jun. 2019, doi: 10.1049/iet-ipr.2018.6212.

[48] G. Mushtaq and F. Siddiqui, "Detection of diabetic retinopathy using deep learning methodology," IOP Conf. Ser. Mater. Sci. Eng., vol. 1070, no. 1, p. 012049, Feb. 2021, doi: 10.1088/1757-899X/1070/1/012049.

[49] A. Bilal, G. Sun, and S. Mazhar, "Survey on recent developments in automatic detection of diabetic retinopathy," J. Fr. Ophtalmol., vol. 44, no. 3, pp. 420–440, Mar. 2021, doi: 10.1016/j.jfo.2020.08.009.

[50] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," IEEE Trans. Neural Networks Learn. Syst., vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[51] H.-C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning.," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1285–98, May 2016, doi: 10.1109/TMI.2016.2528162.

[52] P. Kim, "Convolutional Neural Network," in MATLAB Deep Learning, Berkeley, CA: Apress, 2017, pp. 121–147.

[53] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," Nov. 2015, doi: https://doi.org/10.48550/arXiv.1511.08458.

[54] P. Kumar, R. Kumar, and M. Gupta, "Deep Learning Based Analysis of Ophthalmology: A Systematic Review," EAI Endorsed Trans. Pervasive Heal. Technol., p. 170950, Jul. 2018, doi: 10.4108/eai.10-9-2021.170950.

[55] T. Xiao, L. Liu, K. Li, W. Qin, S. Yu, and Z. Li, "Comparison of Transferred Deep Neural Networks in Ultrasonic Breast Masses Discrimination," Biomed Res. Int., vol. 2018, pp. 1–9, Jun. 2018, doi: 10.1155/2018/4605191.

[56] M. Singh, G. S. Aujla, and R. S. Bali, "A Deep Learning-Based Blockchain Mechanism for Secure Internet of Drones Environment," IEEE Trans. Intell. Transp. Syst., vol. 22, no. 7, pp. 4404–4413, Jul. 2021, doi: 10.1109/TITS.2020.2997469.

[57] A. Sungheetha and R. Sharma R, "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network," J. Trends Comput. Sci. Smart Technol., vol. 3, no. 2, pp. 81–94, Jul. 2021, doi: 10.36548/jtcsst.2021.2.002.

[58] M. A. Habib Raj, M. Al Mamun, and M. F. Faruk, "CNN Based Diabetic Retinopathy Status Prediction Using Fundus Images," in 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 190–193, doi: 10.1109/TENSYMP50017.2020.9230974.

[59] R. H. Paradisa, D. Sarwinda, A. Bustamam, and T. Argyadiva, "Classification of Diabetic Retinopathy through Deep Feature Extraction and Classic Machine Learning Approach," in 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Nov. 2020, pp. 377–381, doi: 10.1109/ICOIACT50329.2020.9332082.

[60] N. Tsiknakis et al., "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," Comput. Biol. Med., vol. 135, p. 104599, 2021, doi: 10.1016/j.compbiomed.2021.104599.

[61] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," Comput. Electron. Agric., vol. 173, p. 105393, Jun. 2020, doi: 10.1016/j.compag.2020.105393.

[62] "KERAS Homepage." https://keras.io/ (accessed Feb. 18, 2023).

[63] Dhiman, P.; Bonkra, A.; Kaur, A.; Gulzar, Y.; Hamid, Y.; Mir, M.S.; Soomro, A.B.; Elwasila, O. Healthcare Trust Evolution with Explainable Artificial Intelligence: Bibliometric Analysis. *Information 2023, Vol. 14, Page 541* 2023, *14*, 541, doi:10.3390/INFO14100541.

[64] Majid, M.; Gulzar, Y.; Ayoub, S.; Khan, F.; Reegu, F.A.; Mir, M.S.; Jaziri, W.; Soomro, A.B. Using Ensemble Learning and Advanced Data Mining Techniques to Improve the Diagnosis of Chronic Kidney Disease. *International Journal of Advanced Computer Science and Applications* 2023, *14*, doi:10.14569/IJACSA.2023.0141050.