



Deep learning-enhanced diabetic retinopathy image classification

Ghadah Alwakid¹, Walaa Gouda², Mamoon Humayun³  and Noor Zaman Jhanjhi⁴ 

DIGITAL HEALTH
Volume 9: 1–15
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076231194942
journals.sagepub.com/home/dhj



Abstract

Objective: Diabetic retinopathy (DR) can sometimes be treated and prevented from causing irreversible vision loss if caught and treated properly. In this work, a deep learning (DL) model is employed to accurately identify all five stages of DR.

Methods: The suggested methodology presents two examples, one with and one without picture augmentation. A balanced dataset meeting the same criteria in both cases is then generated using augmentative methods. The DenseNet-121-rendered model on the Asia Pacific Tele-Ophthalmology Society (APTOS) and dataset for diabetic retinopathy (DDR) datasets performed exceptionally well when compared to other methods for identifying the five stages of DR.

Results: Our propose model achieved the highest test accuracy of 98.36%, top-2 accuracy of 100%, and top-3 accuracy of 100% for the APTOS dataset, and the highest test accuracy of 79.67%, top-2 accuracy of 92.76%, and top-3 accuracy of 98.94% for the DDR dataset. Additional criteria (precision, recall, and F1-score) for gauging the efficacy of the proposed model were established with the help of APTOS and DDR.

Conclusions: It was discovered that feeding a model with higher-quality photographs increased its efficiency and ability for learning, as opposed to both state-of-the-art technology and the other, non-enhanced model.

Keywords

Diabetic retinopathy, vision loss, deep learning, enhanced images, transfer learning, densenet-121, augmentation, DDR, APTOS

Submission date: 21 May 2023; Acceptance date: 28 July 2023

Introduction

The medical field believes that the early discovery of several diseases allows for more effective treatment.^{1–3}

Diabetes is among the most prevalent diseases, and its incidence has risen globally; it is normally related to synthesis of insulin and excessive glucose levels in the organism,^{4,5} leading to metabolic disruption and consequent health issues like cardiovascular disease, kidney failure, mental impairment, and vision loss due to diabetes, among others. Diabetic retinopathy (DR) is a devastating condition that can lead to blindness in advanced stages.^{6–8} The majority of diabetes patients suffer from non-proliferative retinopathy (NPD). In the beginning stages of the disease, the central retina has edema and hard exudates, which are

lipids that possibly leaked from abnormal blood vessels. In the later stages of the disease, the blood flow to the

¹Department of Computer Science, College of Computer and Information Sciences, Jof University, Sakakah, Saudi Arabia

²Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt

³Department of Information Systems, College of Computer and Information Sciences, Jof University, Sakakah, Saudi Arabia

⁴School of Computer Sciences, Taylor's University, Subang Jaya, Malaysia

Corresponding author:

Mamoon Humayun, Department of Information Systems, College of Computer and Information Sciences, Jof University, Sakakah 72341, Al Jof, Saudi Arabia.

Email: mahumayun@ju.edu.sa



retina is cut off due to vascular occlusion, and macular edema gets worse.^{9,10} DR is further classified into NPD and proliferative DR (PD), with NPD being a significantly more severe phase of progressive deterioration as shown in Table 1.^{10,11}

Early detection of DR is challenging since it is asymptomatic or presents with minor indications, leaving a person blind and leading to impaired vision. Consequently, preventing the severity of DR requires prompt diagnosis. This disease's diagnosis demands experts and professionals (ophthalmologists) with pretty efficient technologies and approaches that stimulate breakthroughs in this disease's diagnosis.^{9,12}

Extraction of features using machine learning (ML) approaches was the foundation of the vast bulk of DR research before the problem of manual feature extraction prompted researchers to shift their focus to deep learning (DL).^{13,14}

Subsequent studies in areas of medicine opened the way for the development of numerous computer-assisted technologies, including data mining, image processing, ML, and DL. In recent years, DL has received attention in multiple disciplines, including sentiment classification, handwriting identification, and healthcare imaging analysis,^{15–17} among other fields.⁹

To aid ophthalmologists in their evaluation of possible DR cases, we set out to develop a fast, highly automated, DL-based DR identification. Early detection and treatment of DR can lessen its effects. To accomplish this, we developed a model for diagnostics employing the publicly accessible Asia Pacific Tele-Ophthalmology Society (APTOS)¹⁸ and high-quality dataset for diabetic retinopathy (DDR)¹⁹ datasets utilizing unique images enhancement technique with DenseNet-121²⁰ for model classification.

Furthermore, we highlight the study's contributions.

- The main insight of this research is the use of two filtering algorithms—the contrast-limited adaptive histogram equalization (CLAHE)²¹ and the enhanced super-resolution generative adversarial networks

(ESRGAN)²²—to generate high-quality images for the APTOS and DDR datasets.

- We used augmentation methods to guarantee a steady number of records in both the APTOS and DDR datasets.
- Elements used in a comprehensive evaluation of alternatives to determine the system's efficacy include accuracy (Acc), confusion matrix (CM), specificity, sensitivity, top N accuracy, and the F1-score (F1sc).
- The employed datasets are utilized in a fine-tuning process of the pre-trained network employing the DensNet-121 weight-tuning technique.
- Using a multifaceted training approach supported by a wide variety of training techniques (e.g., data augmentation (DA), batch size, validation patience, and learning rate), the overall dependability of the proposed method is improved, and overfitting is avoided.
- The datasets were used in two stages of model development: training and testing. When the model was put through a rigorous 80:20 hold-out validation, it achieved an impressive 98.7% accuracy in classification when using preprocessing approaches and just 81.23% accuracy when without using enhancement methods for APTOS dataset and 79.6% accuracy in classification when using enhancement approaches and just 79.2% accuracy when without using enhancement methods for DDR dataset.

This research shows 2 scenarios: in scenario 1, an ideal strategy for DR stage enhancement employing CLAHE and ESRGAN techniques, while in scenario 2, no enhancement is employed. In addition, each model's weights were trained using DenseNet-121. Images from the APTOS and DDR datasets have been used to compare the outcomes of the models of the two scenarios. Due to the class imbalance in both datasets, augmentation techniques are required for oversampling. We shall adhere to this plan while we continue writing the paper. Section “Related work” offers background for the consideration of relevant work. Following a discussion of the approach proposed in Section “Research methodology,” the findings are presented and analyzed in Section “Experimentation outcomes.” This investigation is wrapped up in Section “Conclusion” (Table 2).

Related work

When DR picture detection was performed by hand, there were a number of problems. Many people in poor countries have problems because there aren't enough qualified ophthalmologists and examinations are expensive. Automated processing methods have been developed to facilitate early detection, which is critical in the struggle against blindness since it allows for more precise and timely treatment and identification. ML models that were learned on images of the fundus of the eye have lately been able to accurately automate DR identification.

Table 1. Diabetic retinopathy stages.

NPD	Stage 0	NO DR
	Stage 1	Mild DR
	Stage 2	Moderate DR
	Stage 3	Severe DR
PD	Stage 4	Proliferative DR

Table 2. List of abbreviations.

Acronyms	Used for
DR	Diabetic retinopathy
DL	Deep learning
CLAHE	Contrast-limited adaptive histogram equalization
ESRGAN	Enhanced super-resolution generative adversarial networks
APTOS	Asia Pacific Tele-Ophthalmology Society
DDR	High-quality dataset for diabetic retinopathy
NPD	Non-proliferative DR
PD	Proliferative DR
ML	Machine learning
Acc	Accuracy
CM	Confusion matrix
F1sc	F1-score
DA	Data augmentation
ROC	Receiver operating characteristic
PCA	Principal component analysis
CFP	Single-color fundus image
CBAM	Convolutional block attention modules increase discrimination
DNN	Deep neural network
T^p	True positive
T^n	True negatives
F^p	False positive
F^n	False negatives
CNN	Convolutional neural network
LB-CNN	Local binary-convolutional neural network

Automatic methods that collaborate well and don't cost much have taken a significant amount of effort to establish.^{2,23} The latter means that these methods are already better than their old versions in every way. In DR classification studies, there really are two primary schools of

thought conventional, specialist methods and cutting-edge, ML based methods. Below is a deeper look into each of these techniques. For instance, Costa et al.²⁴ employed instance learning to detect lesions in Messidor dataset. In Ref.,²⁵ Wang et al. describes a two-step strategy based on handcrafted features for diagnosing DR and its severity. Using the histogram of orientation gradients, while Leeza et al.²⁶ devised collection of attributes for use in DR identification. Employing Haralick and multiresolution features, Gayathri et al.²⁷ classified DR binary and multiclass. Pires et al.²⁸ gradually upgraded to a larger CNN model, tried out various augmentation techniques, and trained at multiple resolutions, all while employing the APTOS 2019 dataset. The receiver operator curve (ROC curve) for the model validated with the Messidor-2 dataset was 98.2%. In addition, Zhang et al.²⁹ recommended using fundus pictures to detect and evaluate DR. Ensemble learning was utilized to improve the ROC (97.7%), sensitivity (97.5%), and specificity (97.7%) of the CNN models Inception V3, Xception, and InceptionResNetV2. On the other hand, Math et al.³⁰ suggested a learning model for DR prediction. For an AUC of 0.963, a pre-trained CNN estimated DR at the segment level and correctly classified all segment levels. Hemorrhages are an early sign of DR, Maqsood et al.³¹ suggested an innovative 3D convolutional neural network (CNN) model for locating them, one that employs a pre-trained VGG-19 model for extracting features from the segmented hemorrhages. In total, 1509 photos from various databases were employed in the investigations, yielding an average accuracy of 97.71%. Furthermore, Gundluru et al.³² improved feature extraction and classification by using a DL model using Harris hawks optimization and principal component analysis (PCA) for dimensionality reduction. A three-stage process approach is presented by Yasin et al.³³ At first, retinal images are pre-processed after that the hybrid Inception-ResNet architecture was employed to classify the stages of an image's development. Last but not least, DR is graded according to its severity: mild, moderate, severe, or proliferative. While Farag et al.³⁴ proposes an autonomous DL-based severity diagnosis method using a single-color fundus image (CFP), DenseNet169 encodes visual embedding. Convolutional block attention modules (CBAMs) increase discrimination. Finally, cross-entropy loss trains the model on APTOS dataset.

Gangwar and Rav³⁵ built a custom convolutional neural network (CNN) module on top of a pre-trained InceptionResNetv2. Those models were fine-tuned using data from 2 sources: Messidor-1 and APTOS 2019. In tests, the APTOS 2019 dataset achieved 82.18% accuracy, while the Messidor-1 dataset achieved only 72.33%.

While images from the APTOS, Messidor2, and IDRiD databases are combined to form a new dataset established by Raiaan et al.³⁶ All of the photos in the collection undergo preprocessing and are then enhanced using

geometric, photometric, and elastic deformation techniques. Ret-Net-10 is a foundational model that utilizes a categorical cross-entropy loss function to label DR stages throughout its three blocks of convolutional layers and maxpool layers. The Ret-Net-10 model achieved an impressive 98.65% accuracy in our tests.

Furthermore, Saranya et al.³⁷ built an automated model for early DR identification by analyzing images of the retina and looking for red lesions. U-Net architecture is used to semantically partition red lesions, and noise is removed, and local contrast is enhanced during preprocessing. U-Net's Advanced CNN facilitates the pixel-level class labelling necessary for medical segmentation. The model was evaluated using the IDRiD, DIARETDB1, MESSIDOR, and STARE datasets. The proposed identification method achieved 99% specificity, 89% sensitivity, and 95.65% accuracy on the IDRiD dataset. DR severity classification system showed 93.8% specificity, 92.3% sensitivity, and 94% accuracy on the MESSIDOR dataset. Based on GW and other DL

models, Omneya Attallah³⁸ provides a robust and automated CAD tool.

Several researches have been done to identify and classify DR at various stages simultaneously. Traditional image processing and DL both have their place in such picture categorization challenges.^{9,39-44} Investigations on DR identification and diagnostic approaches revealed certain gaps that must be reevaluated. For instance, due to a lack of relevant data, insufficient work has been devoted to developing and training a distinctive DL model. Many researchers, however, have found high dependability values when employing pre-trained models with transfer learning. In the end, almost all of these studies only trained DL models on raw images, not pre-processed ones. This made it hard to scale up the completed classification infrastructure. In the existing research, multiple layers are added to the architecture of models that have already been trained in order to make a lightweight DR detection system. This makes the proposed system more efficient and useful, which is what users want.

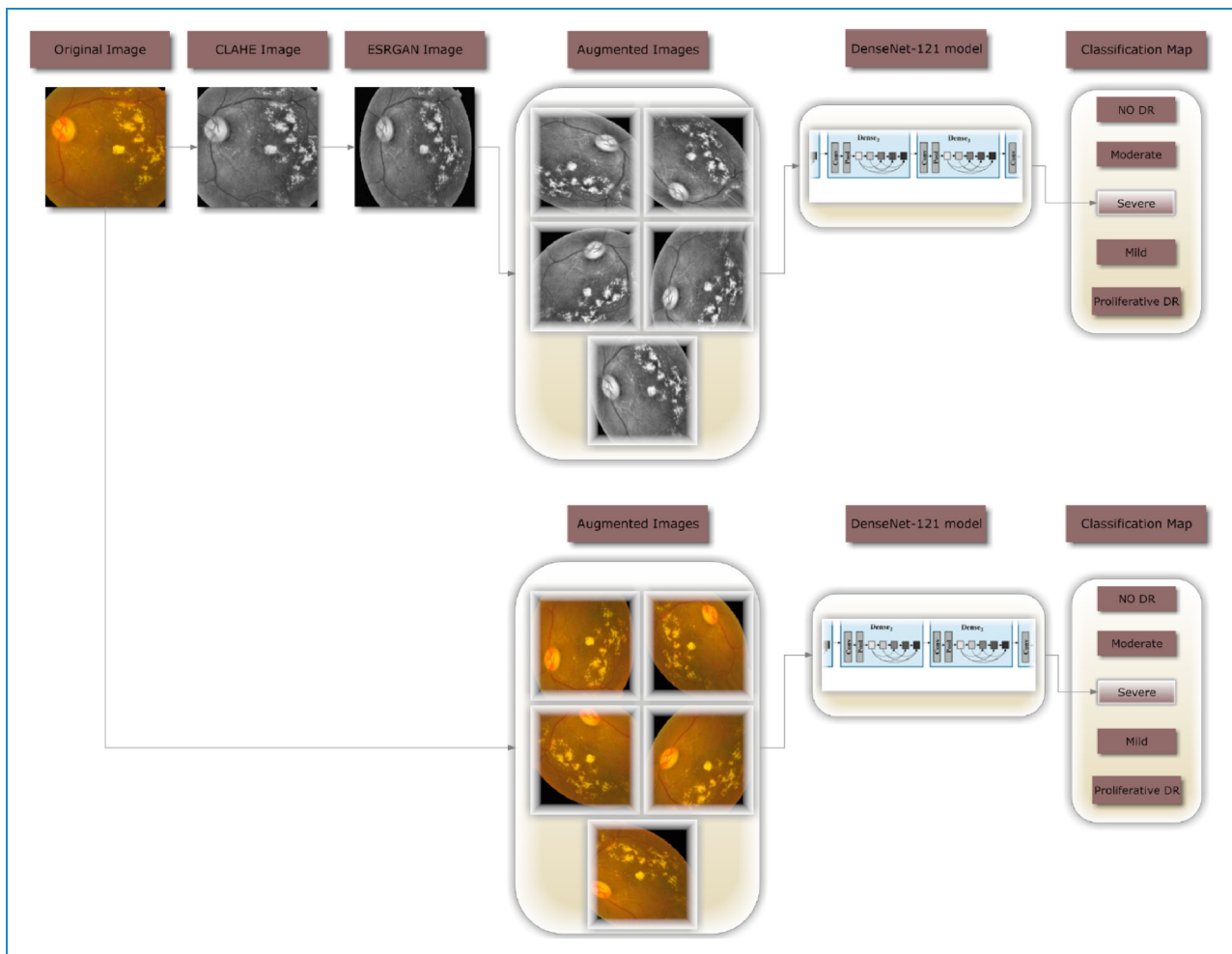


Figure 1. A flowchart for the proposed DR methodology.

Research methodology

In this study, images are classified into one of five DR severity stages using DenseNet-121. Figure 1 illustrates the complete procedure of the proposed methodology that was employed to build a fully automated DR classification model from the dataset discussed in this paper. It illustrates 2 distinct case scenarios: case 1, where CLAHE is utilized as a preprocessing step before ESRGAN, and case 2, where neither of these steps are employed. In both scenarios, the images are augmented to keep them from being too suitable.

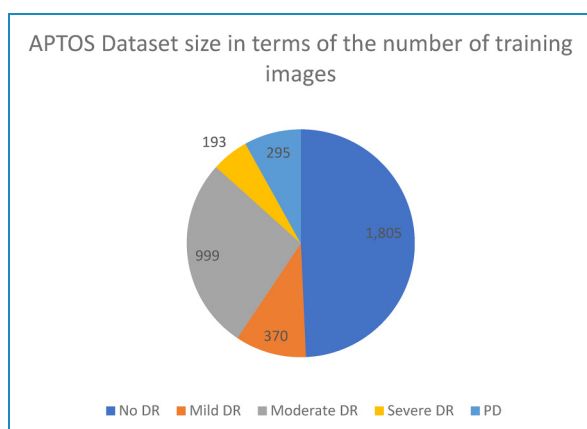


Figure 2. Class-wise image distribution for APTOS dataset.

In the last step, images are sent to the DenseNet-121 model to be classified.

Description of the dataset

APTOS dataset. The amount and quality of images included in the dataset selected are both crucial. The APTOS 2019 Blindness Detection Dataset¹⁸ is used in this study, which is a Kaggle dataset with a large number of photos that is accessible to the world. High-resolution Retinal photos are included in this set, spanning from stage 0 (no DR) to stage 4 (proliferate DR), with designations 1–4 corresponding to the four degrees of severity. Figure 2 displays the distribution of the 3662 retinal images into groups based on the degree of DR present, with 1805 retinal images belonging to the “no DR,” 370 to the “mild DR,” 999 to the “moderate DR,” 193 to the “severe DR,” and 295 to the “proliferate DR.” Several samples of the 3216×2136 pixel image size are shown in Figure 3. Similar to any actual data set, the images and labels contain background noise. There is a chance that the given photographs will contain imperfections, such as blemishes, chromatic aberration, poor brightness, or another difficulty. The photos were acquired throughout time from a wide variety of clinics using different types of equipment, all of which add to the large amount of variance present in the dataset as a whole.

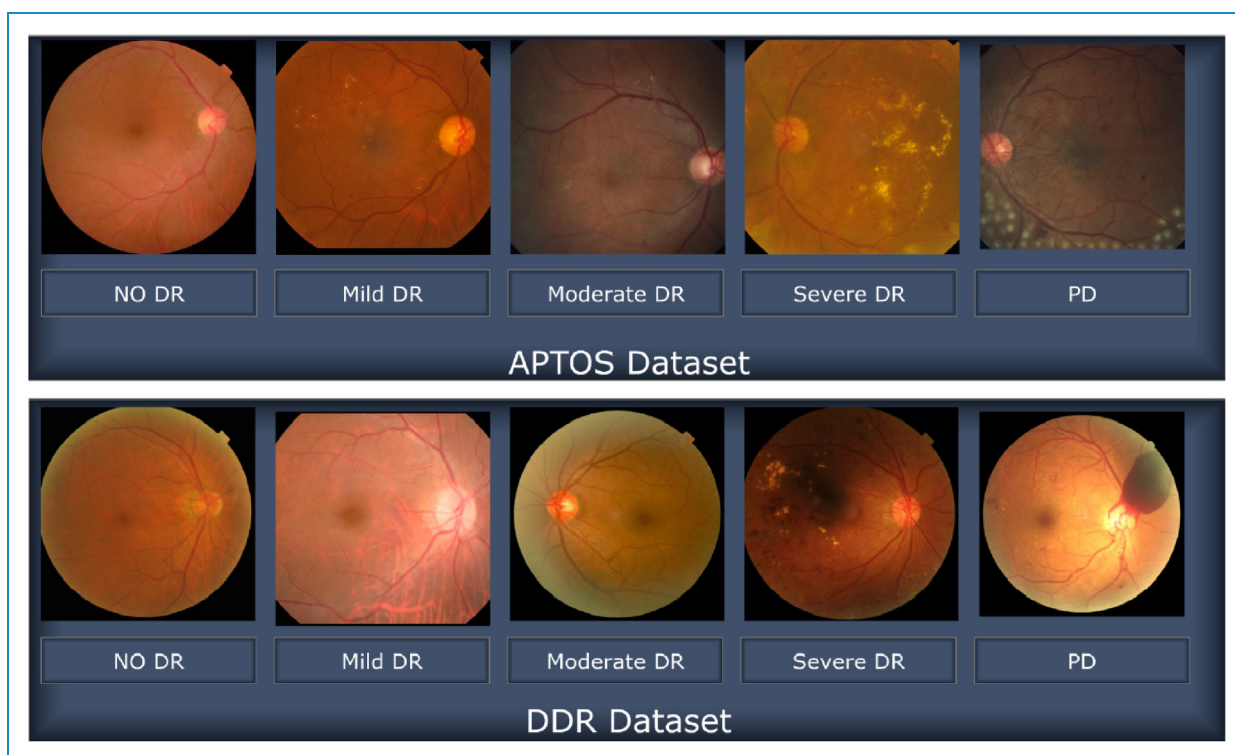


Figure 3. Images representing retinopathy taken from the APTOS and DDR databases.

DDR dataset. This research also utilizes the high-quality DDR dataset,¹⁹ which is a large Kaggle dataset that is available to the general population. The five stages of DR are represented here with high-resolution Retinal pictures, from stage 0 (no DR) to stage 4 (proliferate DR), labelled as stages 0 through 4. The DDR dataset consists of 12,522 images from 147 different universities across 23 different provinces in China (see Figure 4 for distribution). All images in this dataset have had their black backgrounds removed in advance, and the impoverished images from

the sixth category, which represents the low-quality photos, have been omitted as illustrated in Figure 3.

Utilizing CLAHE and ESRGAN for enhancement

Numerous different types of institutions regularly collect photographs of the retina using a wide range of imaging technology. Due to the high contrast enhancement, it was necessary to improve the clarity of the DR pictures used by the suggested method and get rid of various types of noise. Multiple steps are required for all images in scenario 1 to undergo preparatory processing just before augmentation and training. Figure 4b illustrates how CLAHE was used to rearrange the brightness values of the original image to enhance the DR picture's fine features, textures, and low resolution.⁴⁵

This was accomplished by dividing the image into several non-overlapping, nearly identical-sized portions. Consequently, this method simultaneously boosts local contrast and the discernibility of edges and curves throughout an image. Figure 5c shows the second application of ESRGAN on the stage 5 output after all photos have been downsized to the input size of the learning model, which is $224 \times 224 \times 3$. The jagged edges of image artifacts can be mimicked more convincingly in ESRGAN images. When compared to the super-resolution GAN,⁴⁶ the enhanced super-resolution GAN is a significant upgrade.

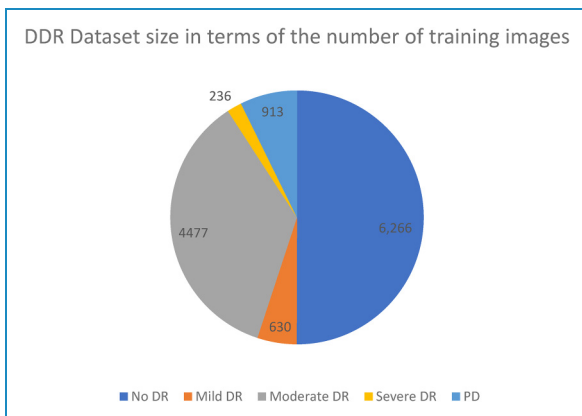


Figure 4. Class-wise image distribution for DDR dataset.

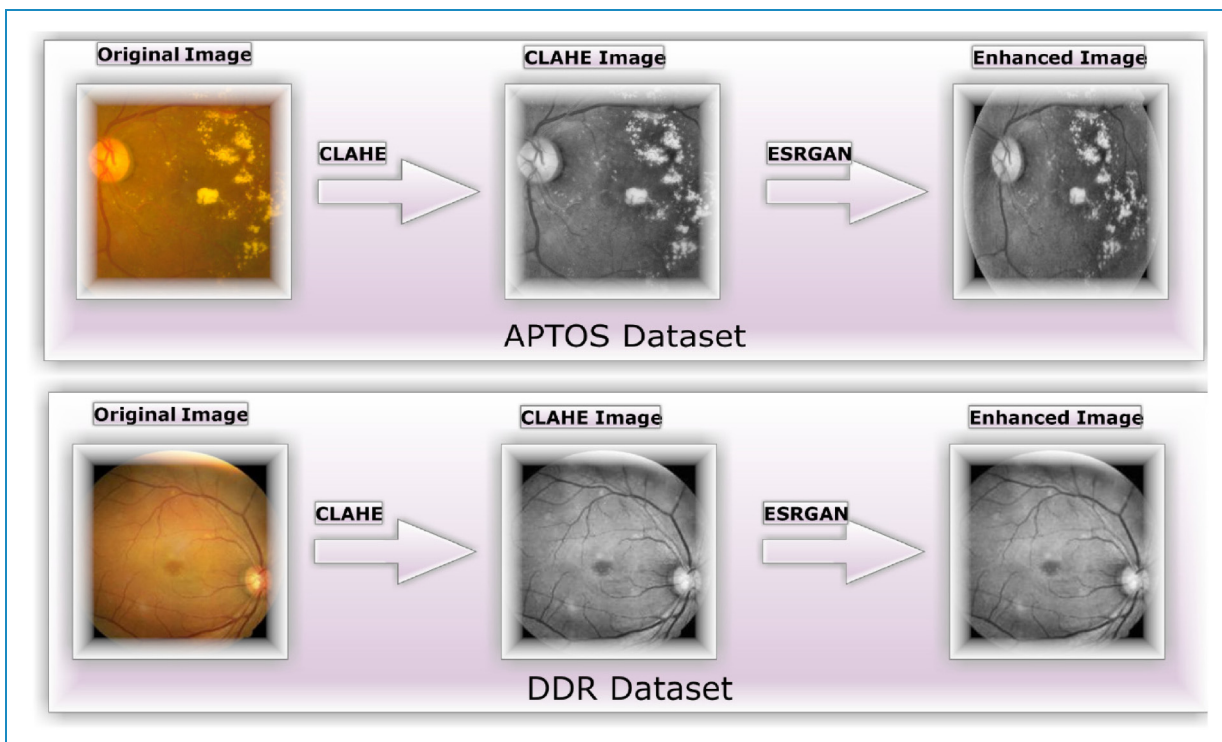


Figure 5. The strategies for improving images include (a) the unaltered original, (b) the identical image produced with CLAHE, and (c) the final product after applying ESRGAN.

Considering the model's minute variations, a residual network is unnecessary. Furthermore, the model lacks a batch normalization layer to smooth the image. Consequently, ESRGAN images can more closely resemble image artifacts with sharp edges. To determine if a picture is real or fake, ESRGAN utilizes a relativistic discriminator.⁴⁷ Using this method yields more precise findings. As a loss function, rival training employs perceptual differences between actual and false images.

Normalizing the intensity of each pixel in an image to a value between [-1] and [1] ensures that the data is uniform and free of noise. Normalization reduces weight sensitivity, making optimization easier. Consequently, the method represented in Figure 5 enhances the quality of the image's boundaries and arches as well as enhancing the image's contrast, resulting in more precise results when employing this technique.

Augmenting data

In order to rectify the existing class divide, it is necessary to raise the overall number of images taken, as depicted in Figures 2 and 4, researchers performed DA on the training set before feeding DenseNet-121 the dataset photos. When given more information, deeper learning models tend to perform better. By making a number of alterations to each photo, we may make use of DR photography's unique qualities. The deep neural network (DNN) retains its accuracy even after the image is enlarged, flipped either vertically or horizontally, or rotated by a predetermined degree of angles. To avoid overfitting and rectify the imbalance in the dataset, DAs (i.e., translation, rotation, and magnification) are used. Among of the modifications examined herein involves horizontal shifting augmentation, which involves moving pixels horizontally while retaining the image's aspect ratio, having the step size specified as an

integer between 0 and 1. The image can be rotated randomly by selecting an angle between -180 and 180 degrees, which is a different kind of transformation than the more frequent width and height translations and zooming in and out. All previous edits to the images within the training set are applied to generate new samples for the network.

In this research, DenseNet-121 was trained using two different scenarios: the first involved applying augmentation to the enhanced photos (shown in Figure 6), and the second involved applying augmentation to the raw images (shown in Figure 7). While the total number of images remains the same, the goal of DA is to enhance the amount of data by adding significantly modified copies of either existing data or newly synthesized data gathered from the available data using the same settings in both cases.

Differential categorization techniques were used to solve the issue of inconsistent sample numbers and unclear groupings. Indicative of a "imbalanced class," where samples are not distributed uniformly across all classes, are the APTOS and DDR datasets (see Figures 2 and 4). In both scenarios, it is clear that the classes are balanced after augmentation procedures have been applied to both datasets as shown in Figure 8.

Experimentation outcomes

DenseNet-121 instructions and setup

To demonstrate the DL system's efficacy and compare outcomes to best practices, tests were carried out on both datasets. Per the proposed training scheme, the dataset was split into three groups: 80% for training (9952 photographs), 10% for test (1012 photos), and a random 10% for validation (1025 photos) to test their capabilities and keep the weight pairings with the best accuracy value during

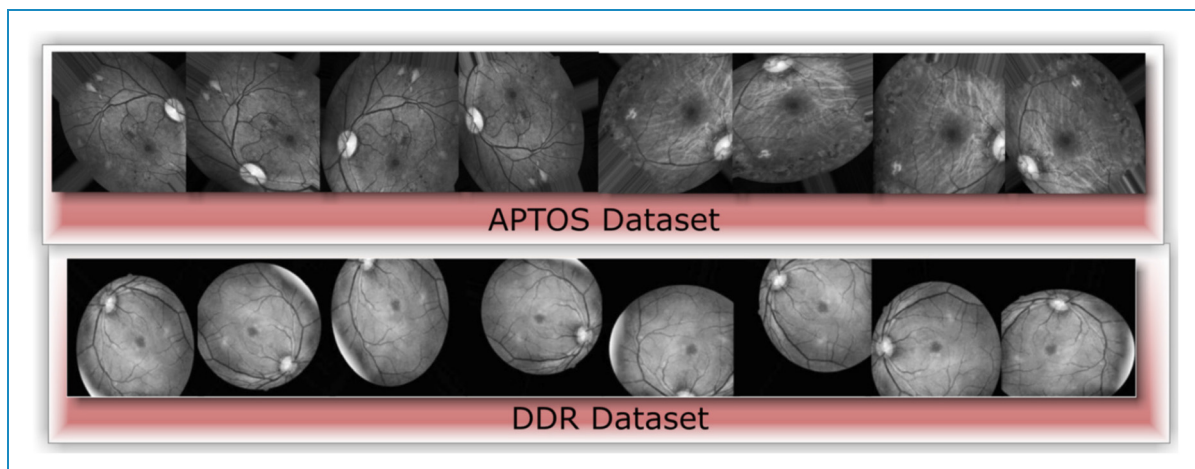


Figure 6. Case 2 scenario DA samples for the APTOS and DDR datasets, using the same image (with improvement).

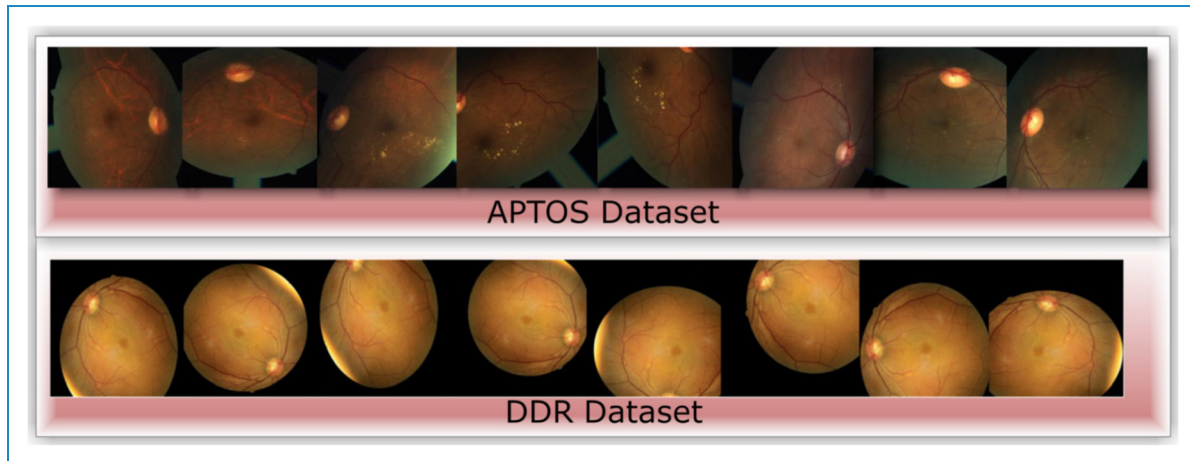


Figure 7. Case 2 scenario DA samples for the APTOS and DDR datasets, using the same image (without improvement).

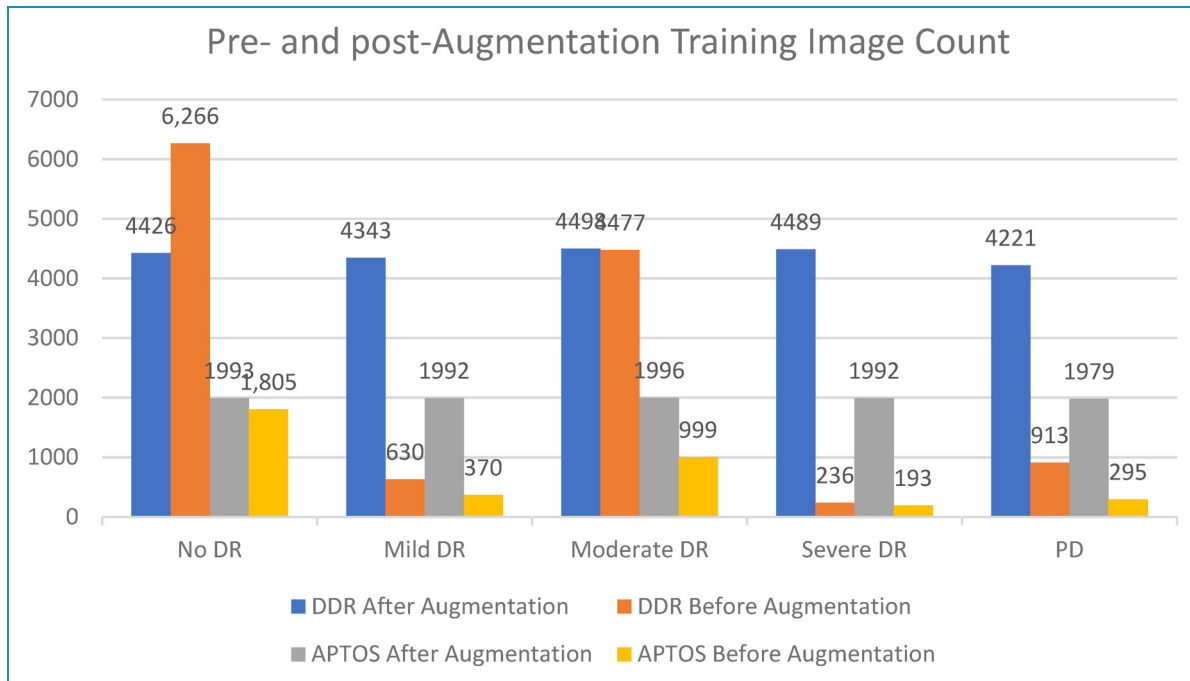


Figure 8. The total number of training photos before and after augmentation case studies on the APTOS and DDR databases.

learning. The training images were shrunk to 224 by 224 by 3 pixels. TensorFlow Keras for the proposed setup was tested on a Linux PC with an RTX3060 GPU and 8GB RAM.

The conceptual approach has already been trained on the datasets using the Adam optimizer, which was given the following hyperparameters: In this experiment, we employ learning rates between $1E^3$ and $1E^5$, batch sizes between 2 and 64 with a 2-fold increase from the previous value, patience equal to 10, epochs equal to 50, and momentum equal to 0.90.

Performance appraisal

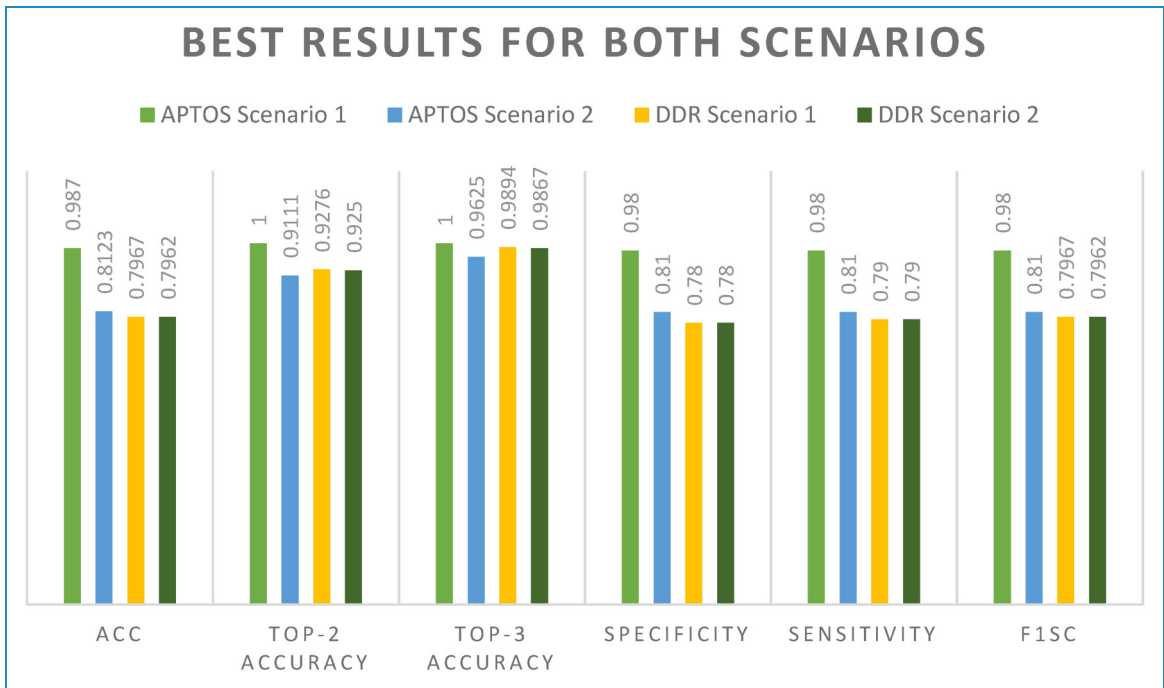
This section of the research describes the evaluation methods used and the results obtained. One popular metric for evaluating classification effectiveness is accuracy (Acc). Formulating this metric requires dividing the total number of examples (images) in the dataset by the number of instances (images) that were correctly labelled, as shown in Equation (1). The two most widely used measures for gauging the efficacy of image classification algorithms are specificity and

Table 3. Top accuracy with enhancement (CLAHE + ESRGAN).

Dataset	Top-2 accuracy	Top-3 accuracy	Acc	Specificity	Sensitivity	F1sc
APTOS	1.0000	1.0000	0.987	0.98	0.98	0.98
DDR	0.9276	0.9894	0.7967	0.78	0.79	0.7967

Table 4. Top accuracy without enhancement (CLAHE + ESRGAN).

Dataset	Top-2 accuracy	Top-3 accuracy	Acc	Specificity	Sensitivity	F1sc
APTOS	0.9111	0.9625	0.8123	0.81	0.81	0.81
DDR	0.9250	0.9867	0.7962	0.78	0.79	0.7962

**Figure 9.** First-rate outcomes in both the APTOS and DDR cases.

sensitivity. The ratio of correctly categorized images in the dataset to those numerically associated is the Sensitivity, and it grows as the number of precisely labelled photos does, as indicated in Equation (3). Since the efficacy of a system cannot be judged solely by its accuracy or sensitivity, a higher F1sc number indicates that the system performs better than one with a lower value. The mathematical derivation of the F1sc is shown in Equation (4). The top N accuracy metric is used at the end of this study, and it measures

how well model N's top replies fit the expected softmax distribution. According to our definition, a categorization is valid if at least one of the N predictions made is in agreement with the desired label.

$$\text{Acc} = \frac{T^p + T^n}{T^p + T^n + F^p + F^n} \quad (1)$$

$$\text{Specificity} = \frac{T^n}{T^n + F^p} \quad (2)$$

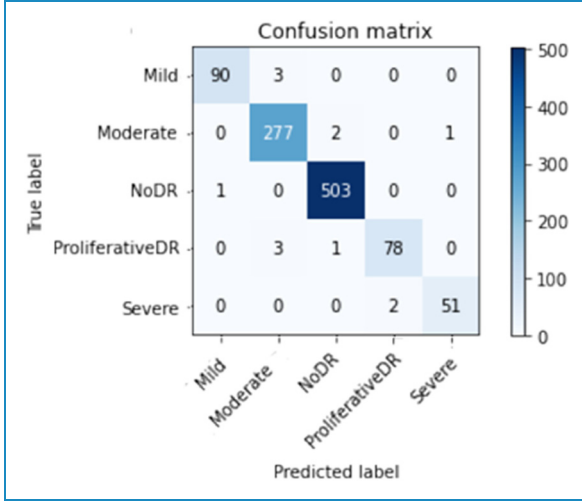


Figure 10. Improved (through CLAHE + ESRGAN) APTOS confusion matrix on top.

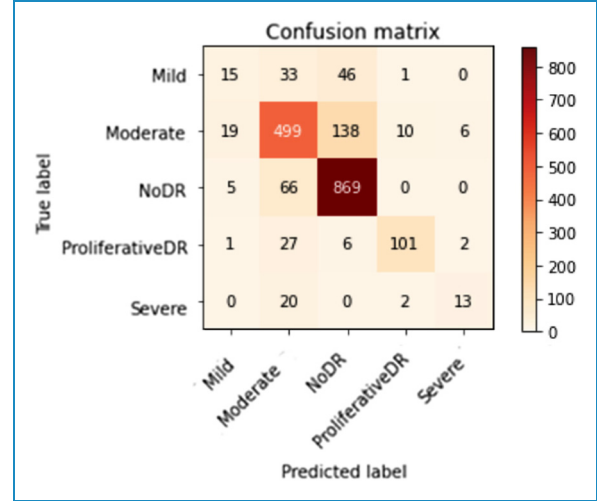


Figure 12. Improved (through CLAHE + ESRGAN) DDR confusion matrix on top.

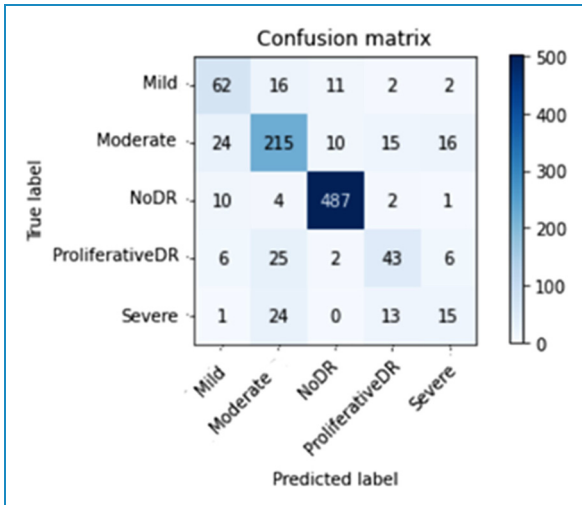


Figure 11. APTOS's top confusion matrix prior to augmentation (i.e., no CLAHE + ESRGAN).

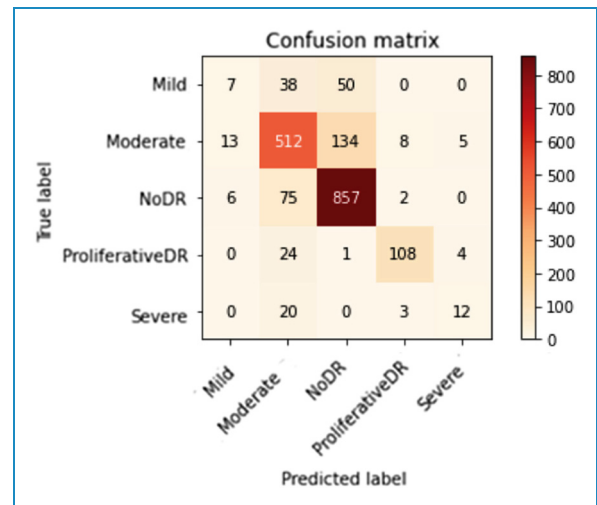


Figure 13. DDS's top confusion matrix prior to augmentation (i.e., no CLAHE + ESRGAN).

$$\text{Sensitivity} = \frac{T^p}{T^p + F^n} \quad (3)$$

$$F1sc = 2 * \left(\frac{\text{Prec} * \text{Re}}{\text{Prec} + \text{Re}} \right) \quad (4)$$

True positives (T^p) and true negatives (T^n) are accurately predicted positive and negative cases, respectively. False positives (F^p) are falsely predicted positive situations, while false negatives (F^n) are wrongly predicted negative situations.

Model results for DenseNet-121's efficacy

When glancing at the datasets, we concentrate on two different scenario configurations in which DenseNet-121 was applied to our datasets in two different ways: once with enhancement (CLAHE + ESRGAN) and then once without (CLAHE + ESRGAN), as shown in Figure 1. Models are trained for 50 iterations, with batch sizes ranging from 2 to 64 and learning rates of $1E^{-3}$, up to $1E^{-5}$. DenseNet-121 was adjusted even more by freezing between 150 and 170 layers. This was done so that it could be as precise as possible. A model ensemble is made up of several runs of the same model with the same

Table 5. Extensive case-by-case analysis for each class.

	Specificity	Sensitivity	F1sc	Total images
CLAHE + ESRGAN for APTOS				
Mild DR	0.95	1.00	0.97	93
Moderate DR	1.00	0.97	0.99	280
No DR	1.00	0.99	0.99	504
PD	0.973	0.98	0.96	82
Severe DR	0.90	0.97	0.93	53
Average	0.98	0.98	0.98	1012
CLAHE + ESRGAN for DDR.				
Mild DR	0.38	0.16	0.22	95
Moderate DR	0.77	0.74	0.76	672
No DR	0.82	0.92	0.87	940
PD	0.89	0.74	0.80	137
Severe DR	0.62	0.37	0.46	35
Average	0.78	0.80	0.78	1879
No CLAHE + ESRGAN for APTOS.				
Mild DR	0.60	0.67	0.63	93
Moderate DR	0.76	0.77	0.76	280
No DR	0.95	0.97	0.96	504
PD	0.57	0.52	0.55	82
Severe DR	0.38	0.28	0.32	53
Average	0.81	0.81	0.81	1012
No CLAHE + ESRGAN for DDR				
Mild DR	0.27	0.07	0.12	95
Moderate DR	0.77	0.76	0.76	672
No DR	0.82	0.91	0.86	940
PD	0.89	0.79	0.84	137
Severe DR	0.57	0.34	0.43	35
Average	0.77	0.80	0.78	1879

Table 6. APTOS dataset efficiency comparison with prior studies.

Ref#	Technique	Accuracy
49	EfficientNet-B6	86.03%
50	SVM	94.5%
51	SVM classifier and MobileNet_V2 for feature extraction	88.80%
52	Densenet-121, Xception, Inception-v3, Resnet-50	85.28%
35	Inception-ResNet-v2	72.33%
53	MobileNet_V2	93.09%
54	EfficientNet and DenseNet	96.32%
55	VGG-16	96.86%
56	Convolutional neural network (CNN)	85%
57	Hybrid Residual U-Net	94%
41	Inception-ResNet-v2	97.0%,
58	VGG-16	74.58%
43	VGG-16	73.26%
	DenseNet-121	96.11%
59	Local binary-convolutional neural network (LB-CNN)	97.41%
4	Inception-v3	88.1%
9	DenseNet201	93.85%
2	MSA-Net	84.6%
Proposed methodology	DenseNet-121 (without using CLAHE + ESRGAN) Scenario 2	81.23%
Proposed methodology	DenseNet-121 (using CLAHE + ESRGAN) Scenario 1	98.7%

characteristics. Since the weights are chosen at random for each run, the accuracy changes from run to run.¹⁵ In Tables 3 and 4, for Scenarios 1 and 2, only the best run result is shown. It shows that the best results you can achieve using CLAHE + ESRGAN in Scenario 1 and Scenario 2 are 98.7% and 81.23%, for APTOS dataset and 79.67% and 79.62 for DDR, respectively. Figure 9 shows how the evaluation metrics used in scenario 1 utilizing CLAHE

Table 7. DDR dataset efficiency comparison with prior studies.

Ref#	Technique	Accuracy
60	NN	66.68%
61	MobileNet + Category Attention Block + Global Attention Block	78.13%
62	Resnet18 (Quasi-Hyperbolic Momentum optimizer)	79.6%
Proposed methodology	DenseNet-121 (without using CLAHE + ESRGAN) Scenario 2	79.62%
Proposed methodology	DenseNet-121 (using CLAHE + ESRGAN) Scenario 1	79.67%

and ESRGAN and scenario 2 without utilizing them affected the best outcome for each scenario.

Figures 10 to 13 depict scenario 1 and scenario 2 confusion matrices. The confusion matrix is a predictive analytic instrument. Regarding machine learning, the confusion matrix is used to measure how well a classification-based machine learning model works.⁴⁸ The confusion matrix shows that the suggested technique employed in scenario 1 can distinguish retina classes with 98.7% accuracy for APTOS and 79.67% for DDR, which is good for real-world application. The confusion matrix demonstrates that scenario 1 correctly classified 503 samples of NODR out of 504 total sample and 869 samples of NODR out of 940 total samples for APTOS and DDR datasets, respectively.

Throughout the datasets, Table 5 displays the full range of test photos by category. The employing of retinal imaging to aid in the diagnosis of eye infections has been demonstrated to be effective in clinical practice.

Assessing several alternative approaches

The method's efficacy is weighed against that of others. Tables 6 and 7 show that in comparison to other options, our approach is superior in terms of both efficacy and performance. In comparison to the state-of-the-art approaches, the suggested DenseNet-121 model improves accuracy to 99.7% for scenario 1 and 81.23 for scenario 2.

Discussion

Relying on CLAHE and ESRGAN, this study came up with a new method for classifying DR. The model that was made was put through its paces using DR images from the APTOS 2019 and DDR datasets. So, there are two training cases scenarios: case 1 scenario with CLAHE + ESRGAN applied to both datasets and case 2 scenario without CLAHE + ESRGAN. For 80:20 hold-out validation for

case 1 and case 2 scenarios, the model had five-class accuracy rates of 98.7% and 81.23% regarding APTOS dataset, respectively, and 79.67% and 79.62% for DDR dataset. For both cases, the proposed method used the DenseNet-121 architecture that had already been trained on the utilized dataset. During the development of the model, we looked at how well it classified two different situations and found that enhancement techniques gave the best results (Figure 8). The general resolution enhancement of CLAHE + ESRGAN is the most crucial component of our method, and we can show that it is responsible for a big improvement in accuracy. Since the DDR dataset contains low-quality images, the general resolution is improved with CLAHE + ESRGAN, but not to the same extent as in the APTOS dataset. The limitation of this study lies in the small sample size. A large enough sample size is essential for valid conclusions to be drawn from a study. Additional samples are needed to enhance the outcomes of testing since a larger sample yields more reliable results.

Conclusion

Researchers have found a way to efficiently and precisely diagnose five different types of cancer by classifying retinal blood vessels in the APTOS dataset. The proposed method utilizes two scenarios: case 1 scenario, which uses CLAHE and ESRGAN to improve the image, and case 2 scenario, which does not use enhancement. Case 1 scenario uses four-step techniques to improve the image's brightness and get rid of noise. Experiments show that CLAHE and ESRGAN are the two stages that have the most effect on accuracy. For the purpose of preventing overfitting and enhancing the general competences of the suggested method, DenseNet-121 was trained on the apex of preprocessed medical images using augmentation techniques. The proposal asserts that when DenseNet-121 is used, the conception model has an accuracy rate of 98.7% for case 1 and 81.23% for case 2 scenarios on the APTOS dataset, and 79.6% in case 1 and 79.2% in case 2 scenarios on the DDR dataset, which is on par with the prediction performance of professional ophthalmologists. The investigation is also unique and important because CLAHE and ESRGAN were used in the preprocessing phase. Research analysis demonstrates that the proposed approach outperforms conventional modeling approaches. For the suggested method to be useful, it needs to be tested on a sizable, complicated dataset that includes numerous future DR cases. Throughout the future, Inception, VGG, ResNet, and other augmentation methods may be used to examine new datasets.

Acknowledgments: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number 223202.

Contributorship: Data curation, Walaa Gouda; Formal analysis, Walaa Gouda and Ghadah Alwakid; Funding acquisition, Ghadah Alwakid; Investigation, Walaa Gouda; Methodology, Mamoona Humayun and Walaa Gouda; Project administration, Mamoona Humayun and Ghadah Alwakid; Supervision, Mamoona Humayun and Walaa Gouda; Writing: original draft, Walaa Gouda; Writing: reviewing and editing, Walaa Gouda and Mamoona Humayun.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval statement: Not applicable.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia (grant number: project number 223202).

Guarantor: Noor Zaman.

Informed consent: Not Applicable.

ORCID iDs: Mamoona Humayun  <https://orcid.org/0000-0001-6339-2257>

Noor Zaman Jhanjhi  <https://orcid.org/0000-0001-8116-4733>

References

1. Association AD. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2014; 37: S81–S90.
2. Al-Antary MT and Arafa Y. Multi-scale attention network for diabetic retinopathy classification. *IEEE Access* 2021; 9: 54190–54200.
3. Hayati M, et al. Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning. *Procedia Comput Sci* 2023; 216: 57–66.
4. Yadav S and Awasthi P. Diabetic retinopathy detection using deep learning and inception-V3 model. *Int Res J Mod Eng Technol Sci*. 2022; 4: 1731–1735.
5. Nall R. An overview of diabetes types and treatments. *Newslett Health Med News Today* 2018; 1: 1–5.
6. Atwany MZ, Sahyoun AH and Yaqub M. Deep learning techniques for diabetic retinopathy classification: a survey. *IEEE Access* 2022; 10: 28642–28655.
7. Pandey SK and Sharma V. World diabetes day 2018: battling the emerging epidemic of diabetic retinopathy. *Indian J Ophthalmol* 2018; 66: 1652.
8. Alwakid G, Gouda W and Humayun M. Deep learning-based prediction of diabetic retinopathy using CLAHE and ESRGAN for enhancement. 2023.
9. Kobat SG, et al. Automated diabetic retinopathy detection using horizontal and vertical patch division-based Pre-trained DenseNET with digital fundus images. *Diagnostics* 2022; 12: 1975.

10. Singh A, et al. Mechanistic insight into oxidative stress-triggered signaling pathways and type 2 diabetes. *Molecules* 2022; 27: 950.
11. Wykoff CC, et al. Risk of blindness among patients with diabetes and newly diagnosed diabetic retinopathy. *Diabetes Care* 2021; 44: 748–756.
12. Alyoubi WL, Shalash WM and Abulkhair MF. Diabetic retinopathy detection through deep learning techniques: a review. *Inform Med Unlocked* 2020; 20: 100377.
13. Amin J, Sharif M and Yasmin M. A review on recent developments for detection of diabetic retinopathy. *Scientifica (Cairo)* 2016; 2016: 6838976.
14. Adak C, et al. Detecting severity of diabetic retinopathy from fundus images using ensembled transformers. arXiv preprint arXiv:2301.00973, 2023.
15. Gouda W, et al. Detection of COVID-19 based on chest X-rays using deep learning. *Healthcare*. 2022; 10: 343. MDPI.
16. Alruwaili M and Gouda W. Automated breast cancer detection models based on transfer learning. *Sensors* 2022; 22: 876.
17. Bajwa A, et al. A prospective study on diabetic retinopathy detection based on modify convolutional neural network using Fundus images at sindh institute of ophthalmology & visual sciences. *Diagnostics* 2023; 13: 393.
18. *APTOS 2019 blindness detection*; 2019, Kaggle: Kaggle.
19. Li T, et al. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. *Inf Sci (Ny)* 2019; 501: 511–522.
20. Huang G, et al. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition 2017.
21. Pizer SM, et al. Adaptive histogram equalization and its variations. *Comput Vis Graph Image Process* 1987; 39: 355–368.
22. Ledig C, et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
23. Gargeya R and Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 2017; 124: 962–969.
24. Costa P, et al. A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images. *IEEE Access* 2018; 6: 18747–18758.
25. Wang J, Bai Y and Xia B. Feasibility of diagnosing both severity and features of diabetic retinopathy in fundus photography. *IEEE Access* 2019; 7: 102589–102597.
26. Leeza M and Farooq H. Detection of severity level of diabetic retinopathy using bag of features model. *IET Comput Vision* 2019; 13: 523–530.
27. Gayathri S, et al. Automated binary and multiclass classification of diabetic retinopathy using haralick and multiresolution features. *IEEE Access* 2020; 8: 57497–57504.
28. Pires R, et al. A data-driven approach to referable diabetic retinopathy detection. *Artif Intell Med* 2019; 96: 93–106.
29. Zhang W, et al. Automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowl Based Syst* 2019; 175: 12–25.
30. Math L and Fatima R. Adaptive machine learning classification for diabetic retinopathy. *Multimed Tools Appl* 2021; 80: 5173–5186.
31. Maqsood S, Damaševičius R and Maskeliūnas R. Hemorrhage detection based on 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients. *Sensors* 2021; 21: 3865.
32. Gundluru N, et al. Enhancement of detection of diabetic retinopathy using Harris hawks optimization with deep learning model. *Comput Intell Neurosci* 2022; 2022: 8512469.
33. Yasin S, et al. Severity grading and early retinopathy lesion detection through hybrid inception-ResNet architecture. *Sensors* 2021; 21: 6933.
34. Farag MM, Fouad M and Abdel-Hamid AT. Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module. *IEEE Access* 2022; 10: 38299–38308.
35. Gangwar AK and Ravi V. Diabetic retinopathy detection using transfer learning and deep learning. in *Evolution in computational intelligence*. Mangalore: NITK Surathkal. 2021; 679–689.
36. Raiaan MAK, et al. A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images. *IEEE Access* 2023; 11: 42361–42388.
37. Saranya P, Pranati R and Patro SS. Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models. *Multimed Tools Appl* 2023; 2023: 1–21.
38. Attallah O. GabROP: gabor wavelets-based CAD for retinopathy of prematurity diagnosis via convolutional neural networks. *Diagnostics* 2023; 13: 171.
39. Majumder S and Ullah MA. Feature extraction from dermoscopy images for melanoma diagnosis. *SN Appl Sci* 2019; 1: 1–11.
40. Majumder S and Ullah MA. A computational approach to pertinent feature extraction for diagnosis of melanoma skin lesion. *Pattern Recognit Image Anal* 2019; 29: 503–514.
41. Crane A and Dastjerdi M. Effect of simulated cataract on the accuracy of an artificial intelligence algorithm in detecting diabetic retinopathy in color fundus photos. *Invest Ophthalmol Visual Sci* 2022; 63: 2100-F0089–2100-F0089.
42. Majumder S and Kehtarnavaz N. Multitasking deep learning model for detection of five stages of diabetic retinopathy. *IEEE Access* 2021; 9: 123220–123230.
43. Yadav S, Awasthi P and Pathak S. Retina image and diabetic retinopathy: a deep learning based approach.
44. Majumder S, Ullah MA and Dhar JP. Melanoma diagnosis from dermoscopy images using artificial neural network. In 2019 5th International Conference on Advances in Electrical Engineering (ICAEE). 2019. IEEE.
45. Reza AM. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J VLSI Signal Process Syst For Signal, Image Video Technol* 2004; 38: 35–44.
46. Wang X, et al. Esrgan: enhanced super-resolution generative adversarial networks. In Proceedings of the European conference on computer vision (ECCV) workshops. 2018.
47. Jolicoeur-Martineau A. The relativistic discriminator: a key element missing from standard GAN. arXiv preprint arXiv:1807.00734, 2018.
48. Townsend JT. Theoretical analysis of an alphabetic confusion matrix. *Percept Psychophys* 1971; 9: 40–50.

49. Maqsood Z and Gupta MK. *Automatic detection of diabetic retinopathy on the edge, in cyber security, privacy and networking*. Bangkok: Springer. 2022; 129–139.
50. Saranya P, et al. Red lesion detection in color fundus images for diabetic retinopathy detection. In *Proceedings of International Conference on Deep Learning, Computing and Intelligence*. 2022. Springer.
51. Lahmar C and Idri A. Deep hybrid architectures for diabetic retinopathy classification. *Comput Meth Biomech Biomed Eng: Imaging & Vis* 2022; 11: 166–184.
52. Oulhadj M, et al. Diabetic retinopathy prediction based on deep learning and deformable registration. *Multimed Tools Appl* 2022; 81: 28709–28727.
53. Lahmar C and Idri A. On the value of deep learning for diagnosing diabetic retinopathy. *Health Technol (Berl)* 2022; 12: 89–105.
54. Canayaz M. Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods. *Appl Soft Comput* 2022; 128: 109462.
55. Escorcia-Gutierrez J, et al. Analysis of pre-trained convolutional neural network models in diabetic retinopathy detection through retinal fundus images. In *International Conference on Computer Information Systems and Industrial Management*. 2022. Springer.
56. Thomas NM and Albert Jerome S. Grading and classification of retinal images for detecting diabetic retinopathy using convolutional neural network. In *Advances in electrical and computer technologies*. Bhillai: Springer. 2022; 607–614.
57. Salluri DK, Sistla V and Kolli VKK. HRUNET: hybrid residual U-net for automatic severity prediction of diabetic retinopathy. *Comput Meth Biomech Biomed Eng: Imaging Vis* 2022; 13: 530–541.
58. Deshpande A and Pardhi J. Automated detection of diabetic retinopathy using VGG-16 architecture. *Int Res J Eng Technol* 2021; 8: 3790–3794.
59. Macsik P, et al. Local Binary CNN for Diabetic Retinopathy Classification on Fundus Images. *Acta Polytech Hung* 2022; 19: 27–45.
60. Rahhal D, et al. Detection and classification of diabetic retinopathy using artificial intelligence algorithms. In *2022 13th International Conference on Information and Communication Systems (ICICS)*. 2022. IEEE.
61. He A, et al. CABNet: category attention block for imbalanced diabetic retinopathy grading. *IEEE Trans Med Imaging* 2020; 40: 143–153.
62. Nanda P and Duraipandian N. A novel optimizer in deep neural network for diabetic retinopathy classification. *Comput Syst Sci Eng* 2022; 43: 1099–1110.