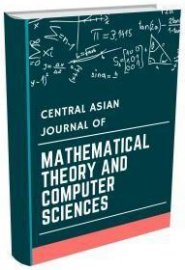


CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.centralasianstudies.org/index.php/CAJMTCS

Volume: 06 Issue: 04 | October 2025 ISSN: 2660-5309

Article

Convolutional Neural Network-Based Classification Models for the Detection of Diabetic Retinopathy in Retinal Fundus

Zainab Fahad Alnaseri

1. University Of Al-Qadisiyah,college of Computer Science and Information Technology

* Correspondence: zainab.alnaseri@qu.edu.iq

Abstract: Diabetes mellitus is a widespread concern worldwide and the most serious microvascular diseases that usually occur as a result are those of the eye, which include the retinopathy and macular edema. Over the past decade, DR has emerged as a significant player in causing vision disability and blindness. Provided that diabetes-related eye complications are promptly diagnosed and handled, the consequences of them may be significantly improved and the level of sugar in the blood can be kept on an adequate level. Nevertheless, the DR symptoms are not consistent and may be complicated; thus, doctors may spend a lot of time to diagnose them.One of the approaches to detecting and classifying DR on fundus retina photographs that is taken into account in the paper is the one which relies on CNNs and deep learning. All the experimental data used in the present study was taken at the Department of Ophthalmology at Xiangya No. 2 Hospital at Changsha in China. The sample of cases is not considerable and the information included in this dataset is imbalanced. That is why a system was made that can be used to rectify the variety and excellence of the information utilized in the training by normalizing and creating information.Then, many CNNs such as "ResNet"-101, "ResNet"-50, and "VGGNet"-16 were employed to ascertain the phases of DR. "ResNet"-101 outperformed the other models by getting 98.88% accuracy and losing 0.3499 during training and 0.9882 during testing. The model was checked on datasets such as HRF, STARE, DIARETDB0, and XHO, which contain 1,787 examples and resulted in an average accuracy of 97%, making it higher than existing methods on the same subject. As a result, using this proposed model enhances DR detection accuracy more than "ResNet"-50 and "VGGNet"-16, making it promising for DR screening in health services.

Keywords: Diabetic retinopathy, classification, Convolutional neural network, deep learning, "ResNet", "VGGNet".

Citation: Alnaseri, Z. F. Convolutional Neural Network-Based Classification Models for the Detection of Diabetic Retinopathy in Retinal Fundus. Central Asian Journal of Mathematical Theory and Computer Sciences 2025, 6(4), 825-845

Received: 10th May 2025
Revised: 16th Jun 2025
Accepted: 24th Jul 2025
Published: 19th Aug 2025

1. Introduction

Both types of diabetes, namely diabetes type 1 and type 2, are said to be factors that give rise to the microvascular disease, known as diabetic retinopathy. Due to such a disease the structure and functioning of the retina become impaired and this happens to be one of the common causes of people becoming blind or losing their vision in the world. According to epidemiologic reports, close to 33 per cent of diabetics are affected by DR and long-term diabetics are almost on the verge of acquiring this condition. It has been estimated that the prevalence of people affected by DR may even touch 191 million by 2030 [1,2].

Blindness caused by diabetes is largely preventable and an early intervention as well as the detection of its impact is beneficial [3,4] Individuals with diabetes that is not controlled well should also undergo screening of DR annually and those that present with the condition should allow the condition to be screened more frequently as the guidelines

outline. During a screening, the quality pictures of the retina are taken and provided to the ophthalmologists to examine thoroughly to make sure that they do not find traces of any retinal damage.

In the absence of treatment, there is an initial development of mild non proliferative stage diabetic retinopathy which tends to progress up to proliferative diabetic retinopathy (PDR). The development of DR is a confirmation of the significance of the usage of the right screening and diagnostic procedures in order to detect the disease in the easiest treatable stage.(Figure 1)

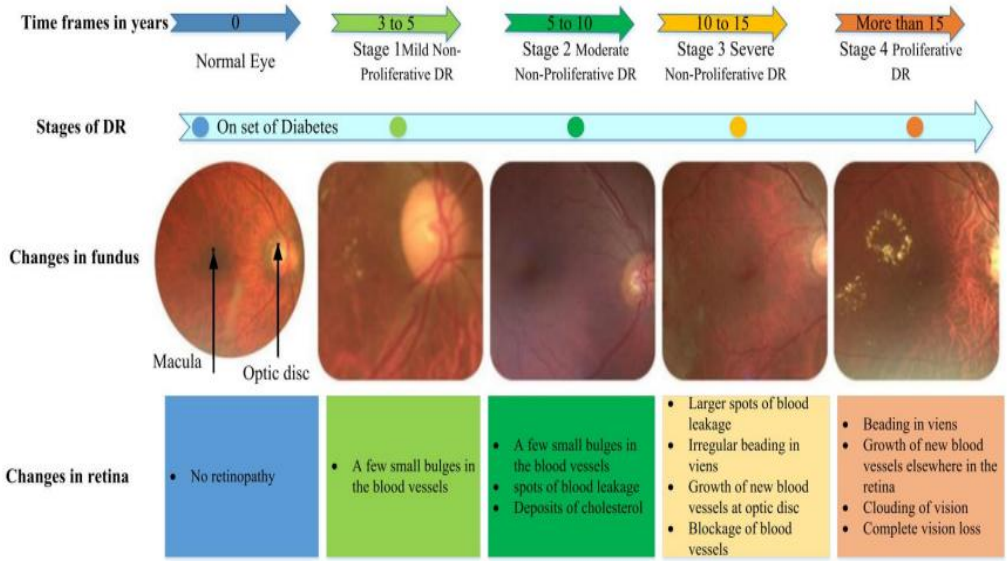


Figure 1. Stages of diabetic retinopathy.

The prevalence of diabetes in the world has led to an expansion in the need of individuals to be screened through the use of an ophthalmologist which poses a significant challenge in the provision of specialized treatment to people. Due to the rise in the number of patients, healthcare systems require enormous quantities of medical personnel and funds to attend to their patients, hence patients take long in seeing an eye doctor [5, 6,7]. Consequently, it becomes evident that the improved automated tools of diagnostics will be useful to ophthalmologists or will serve as autonomous tools to diagnose disabilities in vision. In the recent years, deep learning (DL) methods based on artificial neural networks have contributed to the successful discovery and classification of diabetic retinopathy (DR). Such models are highly sensitive and precise to a great extent as a human being. Also, DL may assist in detecting other eye diseases which are harmful to vision, i.e., diabetic macular edema, glaucoma, and age-related macular degeneration [8], [9], [10],[11].

These tools should be followed by well-established trends in clinical severity in evaluating the retinal images with an aim of assisting in medical care. Although a number of studies have addressed classifying DR in accordance with the past severity scale, there are limited studies which have addressed the classification of maculopathy, owing to the fact that the conditions are unavailable in the data. The problematic aspect with making DL models on retina images is that they require numerous labeled images and the entire process is time-consuming [12], [13], 14]. The focal purpose of the research is an automated method of diagnosing and judging the level of severity of DR through retinal fundus photographs. This is to establish a system that will in real-time correctly classify the clinical stage of DR using fundus photographs. To achieve this objective there are certain image processing techniques employed to enhance the images and select important components which will be employed to perform successful image classification.

" The classification is performed with the help of such models by using CNN architecture as "ResNet"-101, "ResNet- 50, and "VGGNet-16". Both sets of data each model is constructed on are tested and evaluated against training accuracy, training loss, and

testing accuracy. The method assists clinicians to ascertain patients quickly, saves the clinicians time and brings about improved management of diabetic retinopathy in cases of the public patients. The paper recommends a deep learning approach of CNNs to address the issue of diabetic retinopathy (DR) classification. The widespread use of CNN architectures can be explained by many successes of image classification overall and within ImageNet Challenge in particular.

Most health experts apply CNNs due to their effectiveness when it comes to the identification of critical features. Unlike the earlier procedures, CNNs eliminate the procedure of identifying crucial features in a manual way; hence it saves on time and costs. Above all, they do not require much preprocessing, depend less on special features that have to be manually engineered, and discover valuable patterns that exist in images automatically. In addition, CNNs do not simply process the calculations effectively but can succeed in carrying out accurate tasks in recognizing the images. They have greatly contributed to better performances in image recognition using numerous benchmark datasets.

As far as we know, it is the first work in which the classification of DR into symptomatic and asymptomatic types using different CNN architecture, i.e., "ResNet"-101, "ResNet"-50, and "VGGNet"-16, on a proprietary dataset to be collected in association with Xiangya Second Ophthalmic Hospital in Changsha, China, during the March-October 2016 period. The most recent advances of deep learning have also enhanced the adaptability of CNN to huge datasets". This way, we seek a more systematic overview of the power of various CNN models to correctly identify DR and offer a fully automated solution to the problem of bringing the vision impairment burden to a minimal level because of delayed diagnosis.

In comparison to previous research, This study presents improvements in categorization and training effectiveness accuracy for large-scale datasets. The main contributions of this work are summarized as follows:

Preprocessing and Data Enhancement: The dataset in use is high-volume, class-imbalanced, and also inherently noisy, which makes training deep neural networks a problem. To alleviate these factors, a number of preprocessing methods are utilized, such as image resizing, data augmentation, and normalization to enhance the diversity of the datasets and, thus, enhance the overall generalization of models.

State-of-the-Art Performance with CNNs: This paper is published to show that CNNs-based models can also provide state-of-the-art performance in the detection of DR. Critical image features of disease grading are the information learned autonomously by the networks and interpreted and applied to the expert ophthalmologic knowledge. The CNN architectures are rigorously tested and finer points of the extracted features are taken into account focusing on their interpretability and clinical value.

This paper aims at examining the accuracy of three well-known Convolutional Neural Network (CNN) architectures ResNet-101, ResNet-50, and VGGNet-16 to detect and identify the minor visual characteristics related to the various stages of Diabetic Retinopathy (DR). A comparative analysis offers the study the potentials of each of the networks to identify associated fine-grained variations in retinal fundus images. The intended purpose is to identify the best architecture to be deployed practically in the task of medical images analysis, especially in automated classification of Stages of DR.

Criterion to Compare Models and Select Models:

In order to compare the three CNN models, extensive analysis is provided through the major performance indicators as such, classification performance, training performance, and generalization capability. The given comparison helps understand which is the most effective model to classify DR having found the right balance between the predictive performance and the practicality of the model in terms of computing.

Organization of paper: The rest of the paper is organized in the following way(Section 2 provides a comprehensive review of the available literature on DR classification with

CNNs, Section 3 provides the methodology that was used in this study, Section 4 gives the set of experiments, including dataset characteristics, preprocessing, model training processes and an in-depth discussion of the results obtained by each CNN model, Section 5 concludes this paper with a conclusion and a discussion of possible future directions)

Related Work

Initial attempts at the detection of diabetic retinopathy (DR) have involved mainly on the processing of the retinal image into basically anatomical and pathological elements such as blood vessels, microaneurysms, fovea, exudates, hemorrhages, and the optic disc. These elements are looked at in order to evaluate DR severity. To cite just one example, in [15] a multi-facing two-dimensional Gaussian-matched filter was suggested to identify retinal vasculature, and in [16] Sinthanayothin et al. delocalized the optic disc on the basis of high predominance of local average gray variance. Morphological methods were used to locate microaneurysm in the image of a fluorescein angiography by Baudoin et al. [17]. These techniques present a problem of limited applicability to low quality images and do not classify DR severity directly, even though they work well in localization of retinal structures.

Automatic classification of DR has also explored usual processes of machine learning. These methods typically rely on the manual extraction of the features followed by the classification of the images using the customary classifiers. A set of descriptors of different types of retinopathy was designed using modeling with the use of scale-invariant feature transform (SIFT) [18]. The Bag-of-Visual-Words (BoVW) paradigm was used in the research [19,20] to unify low-level data as Speeded-Up Robust Features (SURF) and mid-level description with the help of the semi-soft encoding. This kind of approach has been very efficient as there are areas under the receiver operating characteristic (ROC) curves of 97.8 and 93.5 percent in exudates and red lesions respectively [21], [22].

To enhance representation, Seoud et al. [23,24] presented the probability of lesions as a probability map through the combination of lesion location, size and probability and scored 0.393 on Free-Response Receiver Operating Characteristic (FROC), which is higher than the previous records. SVMs were used to detect DR using features extracted in vascular as well as exudate regions [25], [26], 27], yet they are extremely sensitive to the quality of images and extremely variable to large amounts of data. Other morphological filtering and watershed transformation based methods have been applied as well to detected optic discs [28,29] though these show a reasonable sensitivity and predictive values on testing small datasets.

"In recent years, the advent of deep learning—particularly convolutional neural networks (CNNs)—has revolutionized the field of computer vision, including medical image analysis, CNN-based models have shown superior performance in feature extraction and classification tasks, including blood vessel segmentation and retinal lesion detection" [30,31]. LeNet-5 [32], one of the earliest CNN architectures, was adapted for vascular segmentation. However, early CNN applications suffered from limitations, such as reliance on small and low-quality datasets and the need for expert-defined features.

The next breakthrough of CNN-based image classification was published as AlexNet [33,34], and won the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC). This achieved success, leading to further more sophisticated architectures, such as VGGNet [35,36], GoogleNet [37] and ResNet [38]. Such architectures are among the best to have enhanced classification in many areas. In this respect, whereas in [39], CNN models were applied to find small differences in the DR pictures with an accuracy in classification being 95.68%.

Other researchers like [40], saw the use of two deep learning models which include; a CNN512 based model that was used to run a five class DR classification on APTOS and DDR datasets, with an accuracy of 84.1 and 88.6 respectively; and a model based on YOLOv3 used to localize the presence of a lesion, which achieved a mean average precision (mAP) of 0.216. Also proposed [41] an SVM-based method of classifying DR in

the three classes namely normal, non-proliferative and proliferative based on features extracted in fundus images.

There have also been CNN based network architectures that have been customized to the DR grading. In another example, [42] employed a four-layer deep CNN to classify DR as normal, mild and severe. The multiscale lesion detection done with scale-invariant representational model was shown to be highly generalizable in [43]. A VGG16-based CNN achieved 0.95 sensitivity and 0.75 accuracy, on a 5,000 image validation set in [44], trained on the Kaggle DR dataset [45]. These techniques, although they have a more oriented classification, are mostly inefficient in the localization of the important pathological areas.

To address this limitation, class activation mapping (CAM) was introduced in [46] to generate spatial attention maps using weighted activations after global average pooling [47]. This technique was later adapted for regression tasks in [30], using a modified VGG16 architecture without fully connected layers to enhance lesion localization.

Other recent methods combined deep learning and image processing in a fully automatic DR diagnosis method. An example is given by [48] that suggested a mixed-detection of glaucoma and DR using artificial neural networks and handcrafted features. In [49], many preprocessing operations were used such as contrast-limited adaptive histogram equalization (CLAHE), morphological operations and Canny edge detection to increase the visibility of vessels before being classified.

An integrated framework pathway of automatic classification of diabetic eye disease (DED) was presented in [50], using traditional image processing interventions (such as segmentation, augmentation) along with new CNN architecture to enhance the accuracy of the classification task. A test of 13 enhancement filters (Gaussian, bilateral, Wiener, PDE based and so forth) was undertaken with the Indian Diabetic Retinopathy Image Dataset (IDRiD) in [51]. Improved images were given as input to a CNN model to test and detect the most creative way of preprocessing.

Despite these advances, several persistent challenges hinder progress in automated DR detection. These include:

1. Limited availability of large, high-quality annotated datasets;
2. Image quality degradation due to distortion or blur;
3. Overfitting of deep learning models on small or imbalanced datasets;
4. Computational limitations in training and deploying deep models.

Figure (2) this research points out the main areas where research needs to be done in DR detection with CNNs. Experts are trying to solve these issues by increasing the availability of suitable datasets. coming up with effective ways to process data and designing simple but effective deep learning models for real use in medicine.

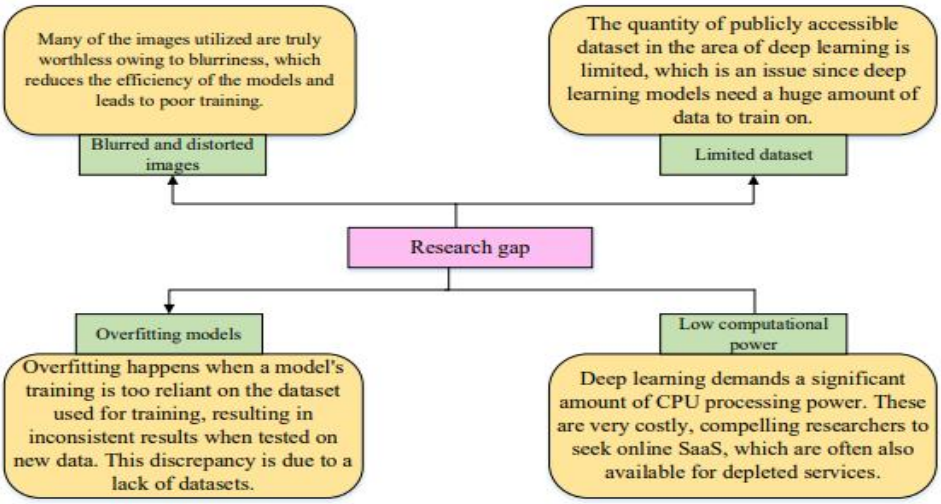


Figure 2. Studying the limitations in the research of Retinopathy by segmenting the lesions with CNN

2. Materials and Methods

Data augmentation and preprocessing were both applied on the laboratory data in different ways to enable deep learning models to perform more efficiently and to be generalizable. These operations were intended to deliver more and diverse training samples and cause the model to cooperate better with the pictures of any quality. Subsequently, procedures known as CNN were applied to process the images and perform a screening of diabetic retinopathy (DR), and the outcomes of multiplying accuracy of their classification were documented.

2.1. Proposed Solution to Address Dataset Challenges

2.1.1. Data Pre-processing

Considering that the quality of images determines the accuracy of a classification to a great extent, the corresponding pre-processing was initiated to bring the data to the standard form and make it more efficient. Images of worse quality may account to errors that decrease the effectiveness of a CNN-based model. Furthermore, the data contained the images of the retina in several such scenarios as the images of different people, age groups, and light levels. Due to these aspects, some pixel intensities differed without linking to the diabetic retinopathy level.

The difficulties were minimized with preprocessing the images with the extensive sequence of functions the OpenCV library offers. Specifically, the geometric transformations were performed via cv2. warpAffine and cv2. warpPerspective, whereas the filters were applied to improve the appearance of an image were denoising ones. The preprocessing involved two main steps only and they were cropping and resizing. 1) Cropping of each fundus image was done by maintaining only circular region of eye and getting rid of unwanted black areas and marks in context with the patient. Due to this method, the significant parcel of the retina remained intact. The images all were cropped, and then they were all put to 300 pixels by 300 pixels. Due to resizing, the data was standardized in terms of dimensions of the input and the accuracy could be learned with regard to the resolution.

"Original Fundus Image Dataset and Retinal Assessment Protocols"

The investigation was conducted based on data of the hospital Xiangya No. 2 Hospital Ophthalmology (XHO), which was nationally known and used experts in diabetic retinopathy monitoring and complications. Retina pictures were captured during March 2016 to October 2016 with the resolution of 1956 x 1934.

All fundus images were classified based on three factors namely (i) presence of diabetic retinopathy (ii) existence of macular edema and (iii) grading capability of the image. The cases were identified as symptomatic or asymptomatic with international scales of diabetes severity of the disease accepted among clinical practitioners. All of them were divided into two categories non-gradeable and gradeable pictures. Only those pictures which could be graded, were taken as ones to use in the model.

The trained ophthalmologists who have more than ten years of experience in grading DR made all the annotations to the images. Each picture was characterized by two ophthalmologists to guarantee quality control. In the event of the diagnosis under debate, the image was not retained in the set so as to ensure the ground truth labels could still be compatible.

Following this rigorous grading process, a total of 1,607 images (607 symptomatic and 1,000 asymptomatic) were allocated for model development and validation. An independent test set comprising 322 images (200 symptomatic and 122 asymptomatic) was used to evaluate model performance, as summarized in Table 1.

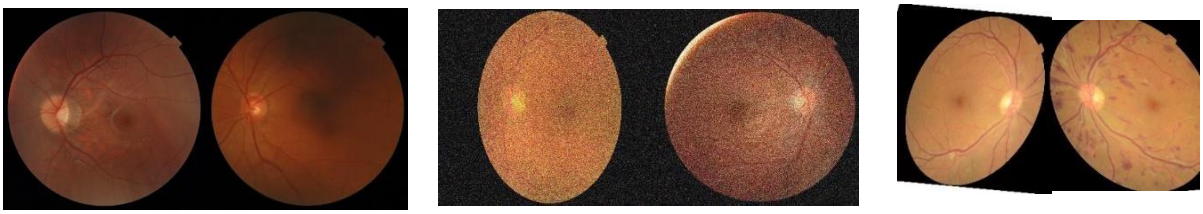
Table 1. dataset for classification.

Class Name	Degree of DR	Number of Labels	For Training	For Testing
Class 0	Nonsymptoms	1000	800	200
Class 1	Symptoms	607	485	122

Training data images (shown in Figure 3) were read and pre-processed into a TensorFlow-GPU library. The image processing type of libraries fits exceptionally well in this library, since it has very rich set of inbuilt functions which have efficient and accelerated computations and therefore when implemented in GPU enabled environments, the computations are more efficient.

One of the most important preprocessing presentation was to reshape the input data to maximize the memory capacity and minimization of its computation. The unprocessed retinal images which measure 1956 x 1934 pixels each require large memory space which might cause more pressure to the RAM and hence slower execution time. In an attempt of mitigating this limitation, all the images were scaled down to an average size of 300 x 300 pixels, thus considerably decreasing the workload without losing important visual characteristics that are central to the classification. Figure 4b and 4c show the consequences of preprocessing and data augmentation actions, respectively.

Figure 3. Examples of retinal picture frames. The two frames in the bottom row are from individuals with diabetic retinopathy, whereas the first two frames in the top row are from healthy participants.



(a) (**b**) (**c**)

Figure 4. Depiction of the preprocessing and enhancement procedures. (a) Authentic photos. Pre-processed photos. (c) Enhanced photos.

2.1.2. Data Augmentation

Deep learning (DL) models are based on lots of data in order to offer effective work. Nevertheless, insubstantiality of data and image noise are some of the reasons which can restrict the learning, and the performance of models. To eliminate these challenges, samples have to be created through data augmentation such that they do not require an increase in the training set through more picture collection. Consequently, the model is more dependable and assists in its avoidance of overfitting.

In the course of the study, the augmentation was conducted with the help of the tools found in Keras. Simple augmentation procedures on the images included rotation of images (in the race to have the model learn to ignore the effect of image rotation), image flipping (horizontally), scaling, clipping, and translating. They allow them to simulate a variety of body shapes and images so as to improve they way the model is fitted to

different patient cases. Also pictures were cut down in such a way that the model concentrated on areas that could contain vital information.

At the recent past, GANs have been used to supplement data since it is able to create realistic images resembling authentic ones. The methodology in [52] demonstrates that it is possible to expand the data in medical imaging using GANs since the information to be trained on can be scarce.

2.1.3. Convolutional Neural Network (CNN) Models

The study is expected to be able to automatically detect and categorize diabetic retinopathy (DR) identified in retinal fundus photographs using CNNs. Due to these special hierarchies sets in CNNs, they can easily learn during the training process, and choose and extract valuable information present in medical images.

The principal components of a CNN structure are convolutional, pooling, and fully connected layers. The convolution layers are responsible to detect local patterns through the usage of numerous optimized filters in the training process. They act similarly to conventional image-processing kernels, though are setup automatically based on data. Max pooling and average pooling assist to minimize data consumed by every layer and also minimize the probability of overfitting due to the retention of critical features. Fully connected layers at the output of the network outline the significant features of the network into classes arranged by the aforesaid categories.

"VGGNet"-16

"VGGNet"-16, a prominent CNN architecture introduced by the Visual Geometry Group, was the winner of the 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC). It consists of 16 layers with learnable weights and approximately 138 million parameters, making it a relatively deep and computationally intensive model.

Despite its large size, "VGGNet"-16 remains a benchmark model for visual tasks due to its simple architecture and consistent performance. In this study, "VGGNet"-16 was employed for DR screening tasks. A modified and compressed version of the original architecture was also utilized, offering the following advantages:

1. **Reduced model size** (88.4% smaller than the original VGG16),
2. **Faster training speed** (23.86% improvement),
3. **Improved convergence and generalization** using residual learning techniques,
4. **Preserved classification accuracy** equivalent to the uncompressed model on large-scale datasets, such as MIT Places 365-Standard.

However, while "VGGNet"-16 demonstrates competitive performance, its classification accuracy was found to be lower than that of deeper architectures such as **"ResNet"-50** and **"ResNet"-101**.

"ResNet" (50 and 101)

Residual Networks ("ResNet"), introduced in [34], revolutionized deep learning by enabling the successful training of ultra-deep CNNs. "ResNet"-152 achieved first place in the 2015 ILSVRC competition, attaining a top-5 classification error of just 3.57%, despite having fewer parameters than "VGGNet".

The core innovation of "ResNet" lies in its **residual learning framework**, which introduces **skip (shortcut) connections**. These connections allow the network to bypass one or more layers, effectively enabling the direct propagation of input features to deeper layers. This mechanism alleviates the **vanishing gradient problem**, ensures **information preservation**, and improves **training convergence**.

In traditional CNNs, increasing depth often leads to a degradation in both training and test performance beyond a certain point. This is not necessarily due to overfitting, but rather a result of optimization difficulties in very deep networks. "ResNet" addresses this degradation by facilitating the learning of **residual functions**, where the network learns the difference F(x)=H(x)−xF(x) = H(x) - xF(x)=H(x)−x, rather than the direct mapping H(x)H(x)H(x). This formulation ensures that identity mappings are easily learned when

additional layers do not contribute positively, thereby enhancing model robustness and accuracy.

In this study, "**ResNet**"**-50** and "**ResNet**"**-101** were selected due to their superior performance in deep feature representation and classification tasks. These models significantly outperformed "VGGNet"-16 in diabetic retinopathy detection by achieving better generalization on the validation set, especially in complex image classification scenarios. (Figure 5)

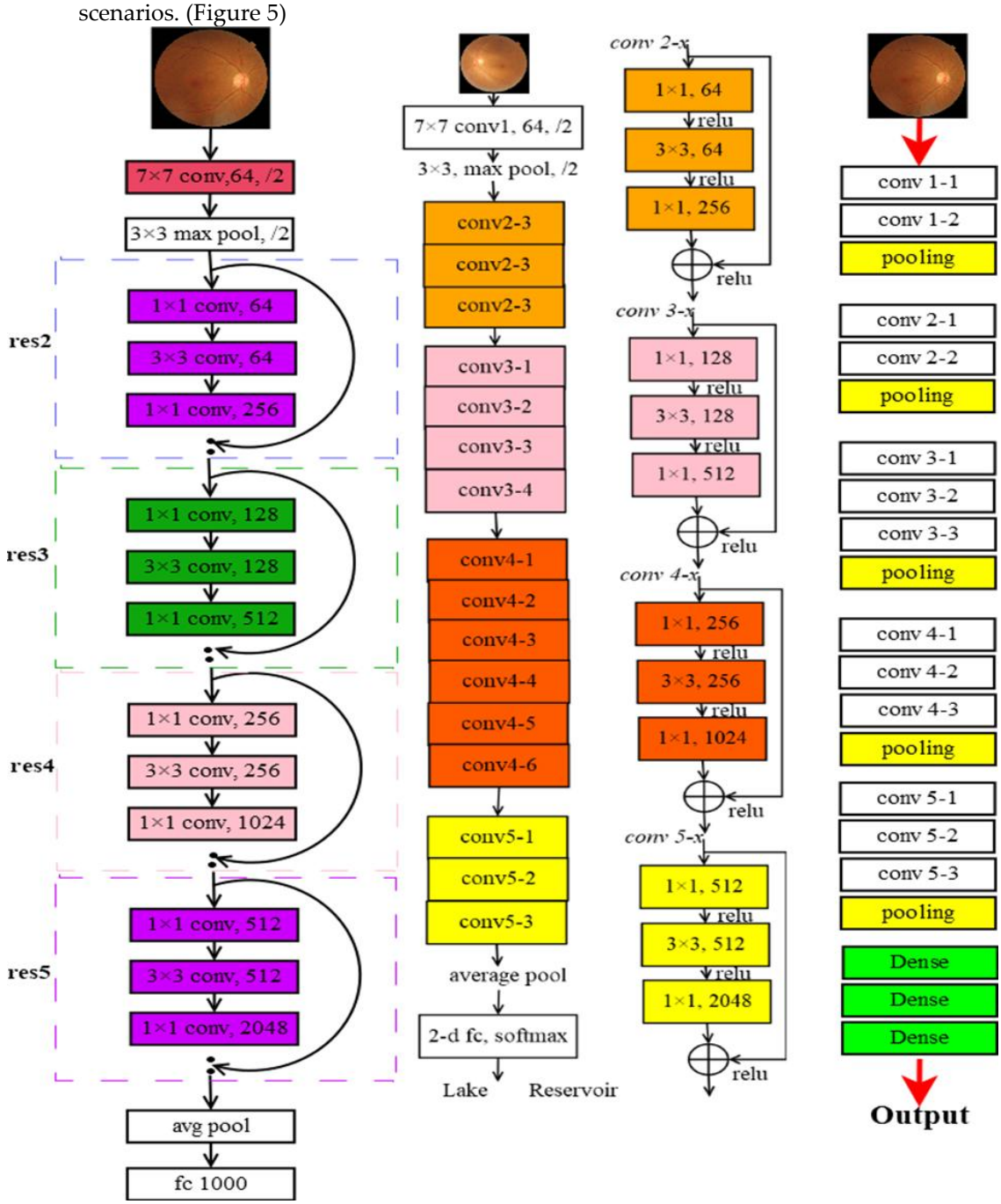


Figure 5. Latest CNN (**a**) "ResNet"-101, (**b**) "ResNet"-V1-50, and (**c**) VGG-16 Architecture.

ResNet-50 is a 50-layers deep convolutional neural network that has been developed to work with the issues that deep network faces once they have become very elaborate. An ImageNet based version of ResNet 50 was trained on more than one million images. That way the feature representations that the model has learned are both robust and varied and can be used to multiple image classification tasks, including medical images.

ResNet -2-101 is also based on ResNet model, but it encompasses a total of 101 layers. This network may be trained on ImageNet where it classified images 1,000 distinct

categories. Due to a detailed structure, the "ResNet"-101 excels in categorizing images, numerous finely-detailed components of which refer to.

In this research, retinal image datasets were evaluated with the help of VGGNet-16, ResNet-50, and ResNet-101 checking whether they had signs of diabetic retinopathy (DR). They have been selected due to their potential use and extensive application in various forms of classification task. In Figures 5a-c, the architectures are depicted; they refer to the architectures of b: ResNet-101, a: ResNet-50 and e: VGGNet-16. Also, the lab dataset demonstrated that DR image analysis involving ResNet -101 is superior to resnet- 50 or VGGNet- 16.

2.1.4. Training Process

To In order to train and test the functionality of the proposed DR framework, the TensorFlow framework in Python was applied with models: ResNet -101, ResNet -V1-50, VGG-16. These models have been chosen due to their favourable performance, good scalability, and ability to accomplish complex work of image classification.

The networks used input images of 300 x 300 pixels since that is what images taken in the retina of the patient were taken to before they were then inputted. The sequence of convolution and pooling layers is followed by placement ReLU layer after each group in order to introduce non-linearity and increase the capability of the network to comprehend.

The aspect of hyperparameter tuning was instrumental in ensuring better models. As opposed to the weights of the model, hyperparameters are not learned; in order to be chosen properly, hyperparameters should be chosen on the basis of an empirical analysis. In this experiment, different hyperparameters were fine-tuned, including learning rate, batch size, the number of epochs, type of optimizer, and a dropout rate, to get the best results by classifying 100 percent of the data and reducing overfitting. Table 2 is a summary of the particular values and settings of the following hyperparameters applied in the implementation of the suggested models.

Table 2. Parameter tuning in the latest CNN models.

Description	Output Shape
weight_decay	0.00001
num_classes	2
batch_size	8
val_batch_size	32
image_height	300
image_width	300
image_channels	3
num_iters	4000
lr	[0.0001, 0.001]
momentum	0.9
batch_norm_scale	True
batch_norm_epsilon	1×10^{-5}
batch_norm_decay	0.997

The data utilized in this research consisted of 1,607 retina images that were collected by a number of hospitals participating in the research in Iraq. e.g., Baghdad, Basra, Mosul, and Dhi Qar. The healthy group had 1,000 images and 607 images revealed presence of DR.

The files were broken into test data and training data so that it became easier to create and analyze the model. Of the whole data we took 800 healthy images and 485 defective pictures to train, and we reserved 200 healthy images and 122 defective pictures to be used in testing. The separation of the cases to normal and pathological components allowed us to compare the results of classification.

They passed through the processing of all pictures and converted them into tensors that can be involved in the convolutional neural networks. The models were then established and refined through the process of continuous training on them and tweaking the hyperparameters. I have gone through the process of tuning the learning rate to begin at 0.0005 by setting the number of iteration in each training of each image to 4,000.

Table 2 contains a comprehensive overview of the hyperparameters used in the training of the created deep learning models.

3. Results and Discussion

Experiment

The study began by implementing a number of pre-processing operations so that the incoming images could be analyzed properly using deep learning. When the training step preconditioned the input images by making them smaller, modifying and normalizing, Deep CNNs performed better on the laboratory acquired images. Thereafter, three highly developed convolutional neural networks (CNNs) were run to extract examples of DR and the accuracy of such networks was calculated. Not only the correct and wrong classifications were counted, but an analysis of errors was particularly studied as well, so that a more rigid analysis could be carried out.

3.1. Dataset and Pre-processing

The research database used in this research was retrieved by four medical institutions in Iraq: Al-Firdaws Private Hospital (Baghdad), Ibn Al-Haytham Teaching Eye Hospital (Baghdad), Al- Shami Eye Center- Erbil Branch and Al-Kafeel Specialist Hospital (Karbala). It had one major aim, which was to differentiate two categories of the patients, those with symptoms and those without symptoms of diabetes. First, the data was heterogeneous and some of the images were not organized, unfiltered and most of them were irrelevant to diabetes. The dataset was then cleaned and tabulated into two clear groups that included the diabetic patients that showed symptoms of retinopathy and the diabetic patients that have not shown these symptoms.

The preprocessing pipeline to which all images were subjected to is aimed at alleviating the problems associated with noise and inadequate illumination. Image resizing and crop were applied in the preprocessing process. Every ouanine was edited to take a focus on a middle circle portion of the fundus that has the most diagnostic features. The size of the original images was rather large (1956 x 1934 pixels), which caused serious difficulties in the computational terms due to the size of the memory that was required. To deal with this, the images were scaled down to 300 x 300 pixels which greatly increased speed of processing and consuming fewer resources.

The total number of pictures that was available initially in the provided dataset was 607 symptomatic and 1000 asymptomatic. Because of the small sample size, it was synthetically oversampled by data augmentation methods. Image rotation and synthetic noise generation with the help of opencv were used as augmentation methods as presented in Table 3 as mythough in Figure 6. Before the augmentation, the training set included 800 asymptomatic images and 485 symptomatic ones and the test set 322 images. There was an increase in size of the training set to 1982 and 1204 images in the asymptomatic and symptomatic class post-augmentation respectively.

The training process was configured with an initial learning rate of 0.001, a momentum value of 0.9, a batch size of 800, and 3000 training iterations.(Table 3)

Table 3. The number of images before and after the augmentation process.

Class	All	Training Set		Tasting Set
		Before Augmentation	After Augmentation	
No symptoms	1000	800	1982	200
Symptoms	607	485	1204	122

3.2. Results and Discussion

In the first experiment, the VGG-16 convolutional neural network architecture (containing 16 layers) was performed to identify the input retinal fundus images. The

architecture uses several layers of convolutions in different scales and alternates with layers of the argmax-pooling operation that decreases spatial sizes, eliminates overfitting, and speeding up the training.

The VGG-16 model was able to successfully separate the training and testing classes as it is stated in the Table 4. The overall accuracy of classification was, but, modest. The main cause of this limitation has been associated with the vanishing gradient problem which makes the model to have difficulties in learning within deep architectures. What is more, the comparatively small depth of network did not seem to allow retrieval of the complex and hierarchical features required to achieve the best diabetic retinopathy indicators classification in the high-resolution retinal photos.

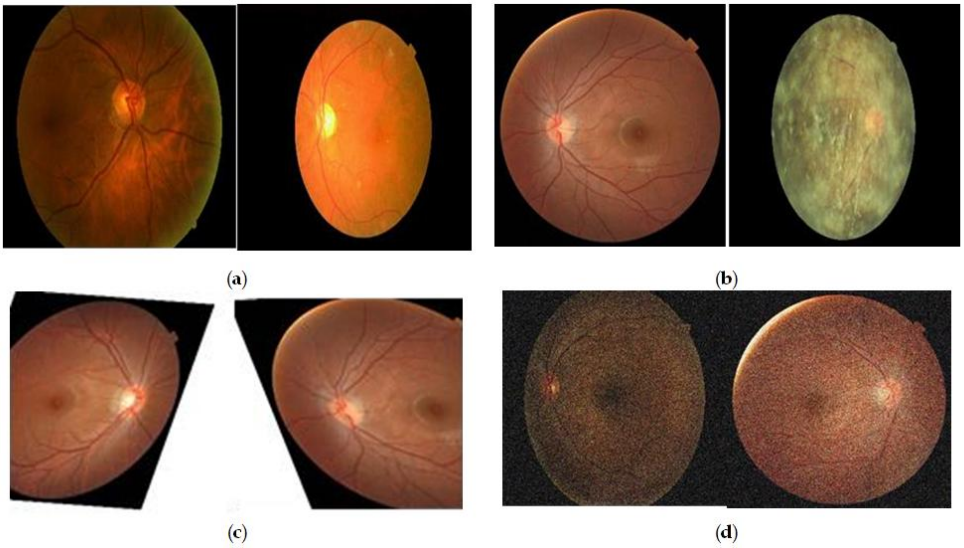


Figure 6. Representative examples from the training dataset illustrating the pre-processing steps employed to augment the input data. (a) Retinal images exhibiting symptoms of diabetic retinopathy. (b) Retinal images without any observable symptoms. (c) Augmented images generated through left and right rotational transformations. (d) Augmented images enhanced with synthetic noise to increase data variability and improve model generalization.

Table 4. Comparative performance evaluation of various classification approaches prior to the application of data augmentation techniques.

Model	Training Accuracy (%)	Training Loss (%)	Testing Accuracy (%)
"ResNet"-101	80.88	40.1	79.76
"ResNet"-50	79.7	37.6	71.8
"VGGNet"-16	64.4	67.32	62.11

In the second experiment, the "ResNet"-50 architecture was used that has 50 convolutional layers with the organizational structure built on residual blocks. Such residual links are particularly crafted in order to overcome the problem of vanishing gradient experienced in deep neural networks. Every residue block has an individual number of convolutional blocks and scales as shown in Figure 5 where the input of a given block is added to the output of the block. This is an architecture design that improves information flow between layers and more abilities of feature extraction. The model ResNet-50 demonstrated better classification performance than VGG-16, and it might be explained by a residual learning framework and greater depth applied in that model.

To continue the increase of the classification performance, the last experiment had a deeper architecture named ResNet-101 comprising 101 convolutional layers. The greater depth enables more abstract and complicated aspects to be learned in the input images to give a better generalization in the model. Although deeper networks are usually prone to overfitting, the residual of the ResNet-101 lessens this chance of overfitting. According to Tables 4 and 5, the accuracy on both the training and testing data proved that ResNet-101 was the most accurate in classification.

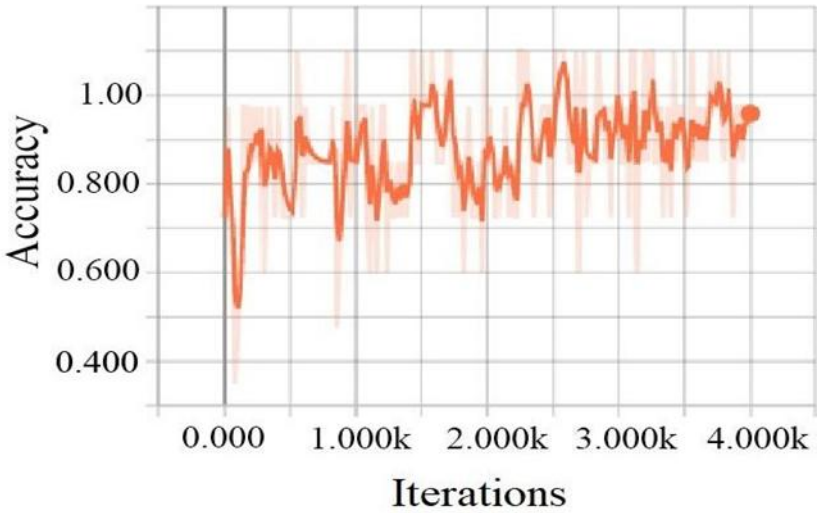
In every experiment, the cross-entropy loss was used in all experiments so that it can direct the search and modify model weights in the training process. The arithmetic mean of the results of the classification was used to assess the model accuracy on the set of testing data.

The results of the CNN models namely, VGG-16, ResNet-50, and ResNet-101 are graphically displayed in Figures 7a 9b". The graphs represent the iterations of the training process (x-axis) versus the values of accuracy and loss (y-axis) of the training process. The performance and stability were most promising in terms of "ResNet"-101 among the three models and showed more potential in recognizing retinal images with better capability of distinguishing between the healthy and diabetic retinopathy-affected retina images".

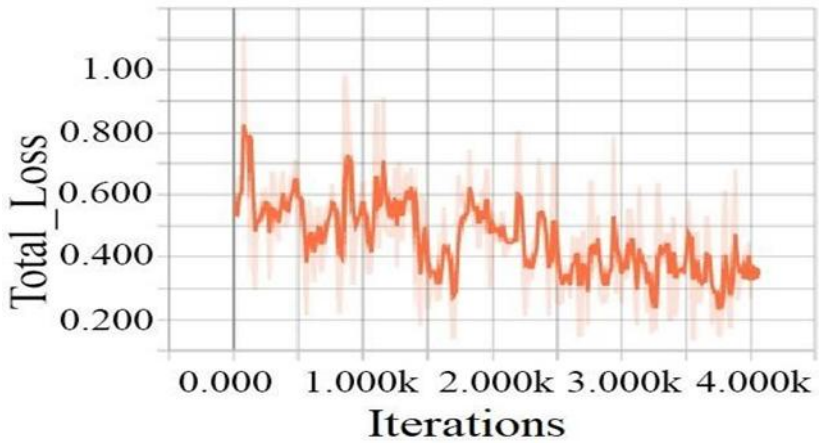
Altogether, the findings show that the offered framework, especially, the model known as the "ResNet"-101, can be theoretical and practical to the classification of images captured by XHO dataset better than the model of one known as the "ResNet"-50 and the one of the VGG-16, in terms of both accuracy and stability.

Table 5. Comparative performance evaluation of various classification approaches following the application of data augmentation techniques.

Model	Training Accuracy (%)	Training Loss (%)	Testing Accuracy (%)
"ResNet"-101	98.88	34.99	98.82
"ResNet"-50	93	34	91.5
"VGGNet"-16	71.39	61.48	64.11



(**a**)



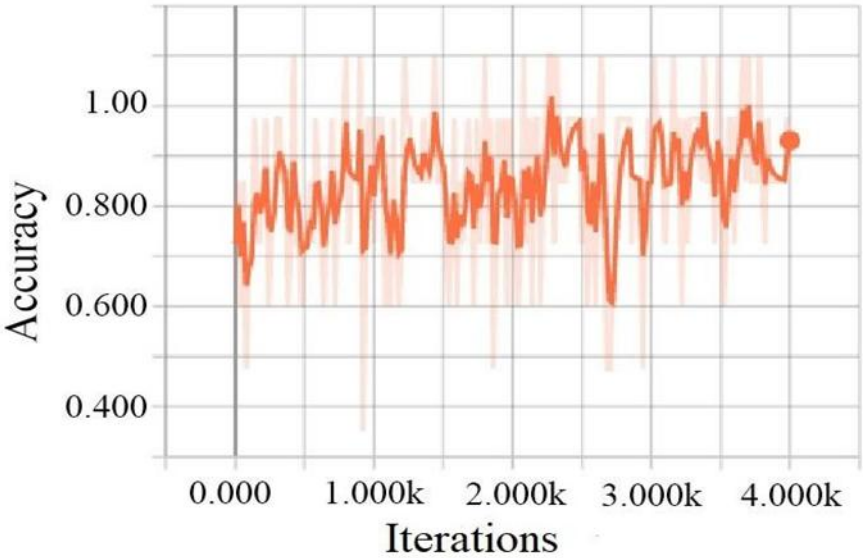
(**b**)

Figure 7. Training performance of the "ResNet"-101 model, illustrating (a) the accuracy progression and (b) the loss reduction over successive training iterations.

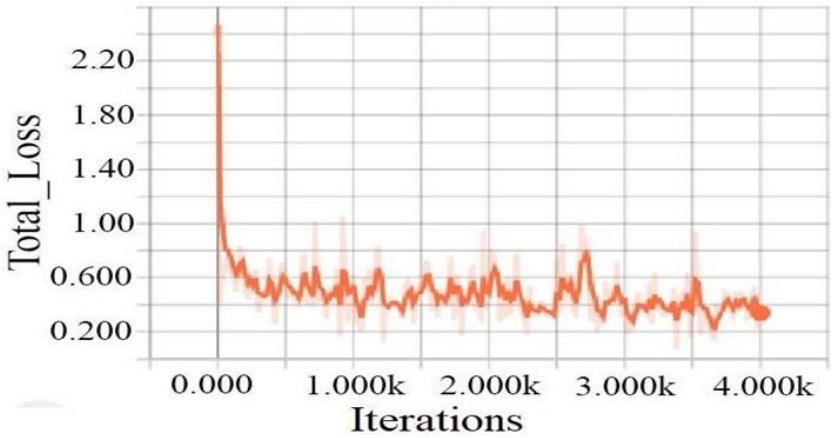
The results demonstrated by classification accuracies shown in Tables 4 and 5 indicate the increase of performance due to data augmentation techniques used. It has been seen that CNN models trained by augmented data perform better than models trained with non-augmented data. This is majorly attributed to the fact that, due to the simulating real-world variation characteristics of data augmentation, this strategy leads to a significant improvement in the robustness of the model, and increases the generalization ability of the model when inferring.

Moreover, the XHO dataset exhibited an imbalanced distribution, with a greater number of normal (non-DR) images compared to images displaying signs of diabetic retinopathy (DR). To mitigate this imbalance, the dataset was initially categorized into two stages of DR and subsequently partitioned into training and testing subsets. Preprocessing techniques were then applied to enhance the quality of the retinal images—an essential step, as poor image quality can significantly impair the feature extraction capabilities of deep learning models and, consequently, reduce classification accuracy. Ensuring consistency in image quality and enhancing salient image features are therefore critical to achieving reliable classification outcomes.

As convolutional neural network architectures and medical imaging datasets continue to evolve, and as real-time diagnostic capabilities become increasingly feasible, the clinical applicability of models such as "ResNet"-101, "ResNet"-50, and VGG-16 for diabetic retinopathy screening is expected to grow. These models have the potential to serve as valuable tools for ophthalmologists, supporting early and accurate detection of DR in clinical practice.(Figure 8)



(a)



(b)

Figure 8. Training performance of the "ResNet"-50 model, illustrating (a) the progression of classification accuracy and (b) the reduction in loss over successive training iterations.

3.3. "Performance Evaluation"

"The proposed diabetic retinopathy (DR) detection approach of the ResNet-101, was tested on various openly available retinal image datasets, besides the in-house XHO data. In particular, the assessment of the approach involved the High-Resolution Fundus (HRF) dataset, comprising 30 images with a resolution of 3304 × 2336 pixels [53]; the Structured Analysis of the Retina (STARE) dataset, containing 20 images of size 700 × 605 pixels [54]; the DIARETDB0 dataset, consisting of 130 images with dimensions of 1500 × 1152 pixels [55]; and the MESSIDOR dataset, which includes 1200 images at a resolution of 1440 × 960 pixels" [56]. The XHO dataset used in this study comprises 1607 retinal images standardized to 300 × 300 pixels.

For evaluation purposes, it was decided to group the testing datasets into those with no DR signs and those that were confirmed as having DR by expert ophthalmologists. Among the overall 2987 images labeled as DR, 1089 were used in the testing phase. The proposed method was tested using normal retinal images to check if it can specifically identify cases without DR.(Figure 9)

Table 6 summarizes all the datasets used in the study and mentions the number of images, their image resolutions, and the distribution of the classes.

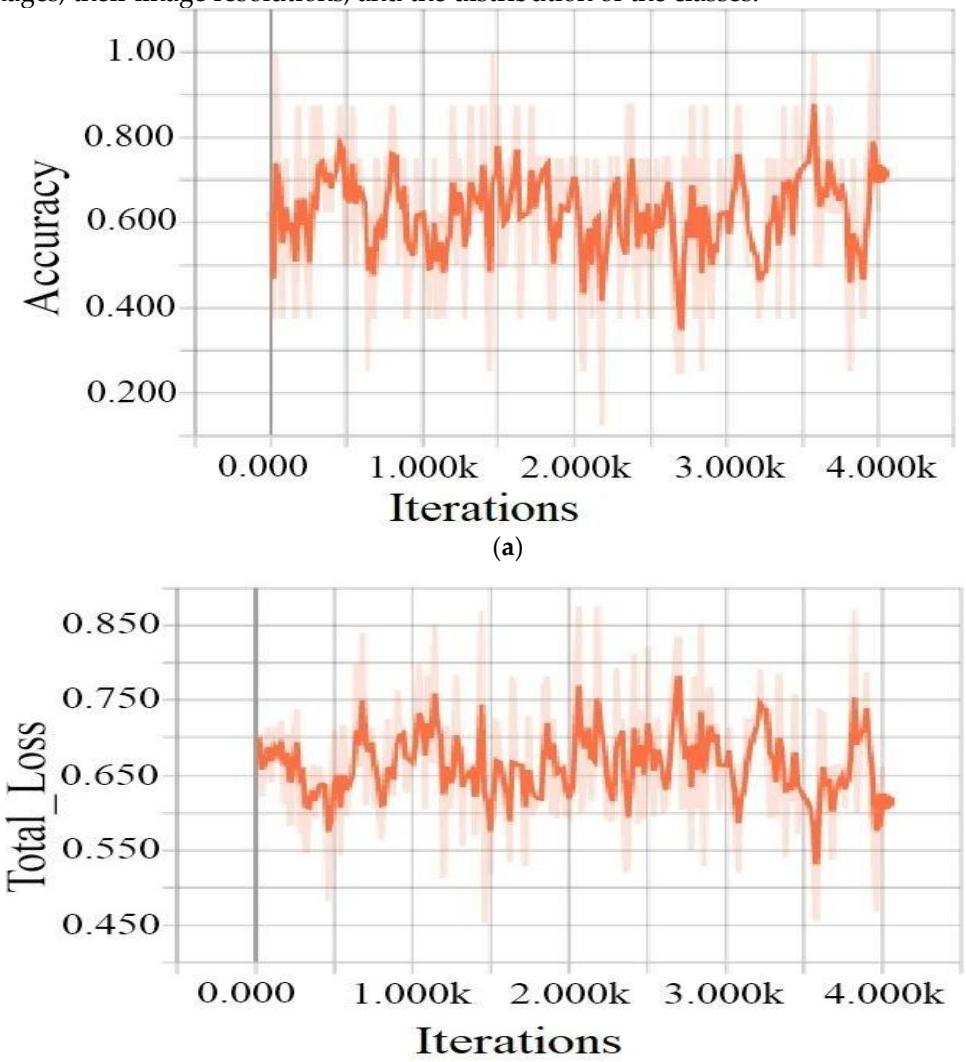


Figure 9. Both the accuracy and the loss curves for "VGGNet"-16 as it is trained.

Table 6. Details of everything in the database.

Dataset	Number of Images	No DR	DR
XHO datasets	1607	1000	607
HRF datasets	30	15	15
STARE datasets	20	12	8
DIARETDB0 datasets	130	20	110
MESSIDOR	1200	851	349
Total	2987	1898	1089

3.4. Performance Evaluation Criteria

The chosen models of neural networks were implemented and their performances were then measured against a pack of steady postulated performance measures that are mostly used in medical image classification systems. These were precision (ACC), specificity (SP), sensitivity (SEN), as well as area under the receiver operating characteristic curve (AUC). These (key) indicators were used to compute the positive predictive value (PPV) to evaluate the model accuracy in the true positive cases. Besides, two more evaluation metrics were used, including negative predictive value (NPV) and the F1 score (F1). All these metrics give a hint of various performance of the model studied. Accuracy (ACC) estimates the percentage of the right classification of the images compared with the number of samples, providing an overall estimate of the success of classification. Specificity (SP) is the one that assesses how well the model discriminates non-diseased (non-DR) cases, thus, its ability to achieve low false positives. Sensitivity (SEN) or recall will measure the proportion of the number of true levels (the number of actual DR cases that were actually positive) that the model properly classifies into the positive category or the true positive, which reflects the performance of the model to identify positive cases of diabetic retinopathy. The AUC gives an overview of the diagnostic capacity of the model at all levels of the classification range. In the meantime, PPV and NPV give the reflection of trustworthiness of affirmative and negative prognoses respectively. F1 score, which is the function of both precision and recall, is especially helpful in cases when the classes are unevenly distributed since it takes into account such tradeoff between false negatives and false positives.

ROC curve eases the comprehension of the way of balancing sensitivity and specificity during algorithm assessment. The AUC can assist us in describing how well the model may discriminate. The PPV is a proportion of correctly labeled DR-positive images out of the number of images that were granted as DR. In contrast, NPV represents the sum of the number of non-DR images, which have been correctly identified as non-DR out of the entire numbers that are predicted to be non-DR.All of these metrics join to help compare the diagnostic results of the "ResNet"-101 algorithm to other algorithms. Music pieces are measured using these formulas in mathematics.

$$SP = \frac{TN}{TN+FP} \quad (1)$$

$$SEN = \frac{TP}{TP+FN} \quad (2)$$

$$ACC = \frac{(TN+TP)}{(TN+TP+FN+FP)} \quad (3)$$

$$PPV = \frac{TP}{TP+FP} \quad (4)$$

$$NPV = \frac{TN}{TN+FN} \quad (5)$$

$$F1score = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

Those diabetic retinopathy (DR) images that are missing a diagnosis of DR are called false negative (FN). to put it another way, FPs are errors where healthy images get labeled as having DR. Correctly identified DR and non-DR images are called true positive (TP) and true negative (TN).

"ResNet"-101 was checked with several measures, among them accuracy (ACC), sensitivity (SEN), specificity (SP), and the F1 score. Two-thousand nine hundred and eighty-seven fundus retinal photos taken from various public sources were used to assess

the model's ability to detect DR. As shown in Table 7, the "ResNet"-101 model's evaluation led to the following ACC, SP, SEN, area under the ROC curve, and F1 score results.

"For the "HRF, DRIVE, STARE, MESSIDOR, DIARETDB0, and DIARETDB1" datasets, "AUC" values were measured, giving useful information about how well the model works on various images". The graphs in Figure 10 clearly show how sensitivity and specificity differ for each of the datasets.(Figure 10)

Results from the "ResNet"-101-based system were compared to those from recent top-performing approaches to see how well it works. As demonstrated in Table 8, "ResNet"-101 does a better job at finding DR than the other two models and is likely a reliable diagnostic method.

Table 7. Restated, the "ResNet"-101 CNN model was used to detect DR from four different datasets.

Dataset	Test Image	Correctly Detected	Accuracy (%) (%)	Sensitivity (%)	Specificity (%)	F1 Score (%)	AUC
Al-Firdaws Hospital (Baghdad) dataset	200	196	98	97.14	97.65	97.36	98.55
Ibn Al-Haytham Teaching Eye Hospital (Baghdad) datasets	30	30	100	99.98	99.98	99.98	99.99
Al-Shami Eye Center - Erbil branch datasets	20	19	95	94.96	95.11	95.03	95.04
Al-Kafeel Specialist Hospital (Karbala) datasets	110	105	95.45	95.39	99.38	95.45	95.46
Westeye Eye Hospital in (Sulaymaniyah) datasets	349	347	99.42	99.45	99.38	99.41	99.42
Total	360	349	97	96.87	98.03	96.95	97.26

Table 8. Looking at the similar techniques applied to DR Images by others and what we have done.

Evaluated Parameter		[57]	[58]	[59]	Proposed Method	[60]	[39]	[61]	[62]	[63]
Number of Classes		2				5		4		5
Detect Lesion		No				No		No		Yes
Dataset		private dataset				Kaggle		Messidor		DDR
Performance	ACC	94.23 %	88.21%	98.7%	98.88%	63.23%	95.6%	98.15%	96.35%	82.84 %
Measure	AUC	0.9823	0.946	-	98.55%	-	0.978	-	-	-

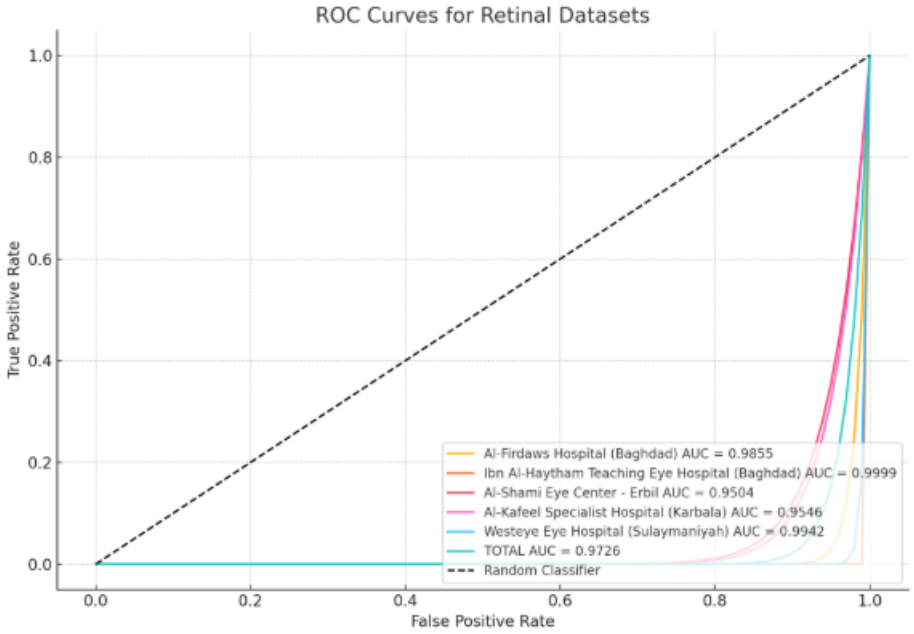


Figure 10. ROC plot for retinal

4. Conclusion

Automatic identification of diabetic retinopathy (DR) within fundus images could go a long way in helping ophthalmologists to greatly reduce the time of diagnosis as well as improving early diagnosis. Here, the recovery of preprocessing and regularizing of the retinal datasets, both of which were received in the laboratory, were done in a comprehensive way in a bid to ensure the deep learning-based classification system is optimized in performance. It was found during initial inquiry that the small training dataset size negatives the model performances and qualifies the requirements of sufficient data diversity and volumes.

Besides, this study proved that DR classification can be successfully tackled as binary classification task to screen the population at the national level using convolutional neural networks (CNNs). One of the considered structures, namely, ".ResNet"-101, ".ResNet"-50 and VGG-16 had the best efficacy. The accuracy of the model in testing was 98.82 percent with corresponding training accuracy of 98.88 percent and a training loss of 34.99 percent".

Upon testing on five benchmark datasets "HRF, STARE, DIARETDB0, MESSIDOR, and XHO", ResNet-101 achieved a classification accuracy of 98%, 100%, 95%,95.45 and 97 respectively. Such findings indicate that convolutional neural network models can be trained to capture the distinguishing characteristics of DR using fundus images and that advanced CNN can train the features of DR in fundus images.

Balance of the quality and classes of datasets is a critical factor in building powerful DR detection systems. Future research ought to concentrate on synthesizing various datasets to create a more capable training set so as to combat the imbalance in data. Although the existing CNNs models perform very well when applied to binary classification, these models cannot cope with the fine-grained classification, or the multi-stages of DR severity. As such, advances will be based on extending these architectures by adding more layers and new designs of architectures together with the ability to classify in real-time to allow more accurate and scalable detection of DR in clinical situations.

REFERENCES

[1] R. Reguant, S. Brunak, and S. Saha, "Understanding inherent image features in CNN-based assessment of diabetic retinopathy," Sci. Rep., vol. 11, p. 9704, 2021.

[2] F. Aguirre, A. Brown, N. H. Cho, G. Dahlquist, S. Dodd, T. Dunning, M. Hirst, C. Hwang, D. Magliano, C. Patterson, et al., IDF Diabetes Atlas, 6th ed. Basel, Switzerland: Int. Diabetes Fed., 2013.

[3] D. S. Ting, K. A. Tan, V. Phua, G. S. Tan, C. W. Wong, and T. Y. Wong, “Biomarkers of diabetic retinopathy,” Curr. Diabetes Rep., vol. 16, p. 125, 2016.

[4] R. R. Bourne, G. A. Stevens, R. A. White, J. L. Smith, S. R. Flaxman, H. Price, J. B. Jonas, J. Keeffe, J. Leasher, K. Naidoo, et al., “Causes of vision loss worldwide, 1990–2010: A systematic analysis,” Lancet Glob. Health, vol. 1, no. 6, pp. e339–e349, 2013.

[5] N. Cheung, P. Mitchell, and T. Y. Wong, “Diabetic retinopathy,” Lancet, vol. 376, pp. 124–136, 2010.

[6] K. Kroenke, “Telemedicine screening for eye disease,” JAMA, vol. 313, pp. 1666–1667, 2015.

[7] DCCT/EDIC Research Group, D. M. Nathan, I. Bebu, D. Hainsworth, R. Klein, W. Tamborlane, G. Lorenzi, R. Gubitosi-Klug, and J. M. Lachin, “Frequency of evidence-based screening for retinopathy in type 1 diabetes,” N. Engl. J. Med., vol. 376, pp. 1507–1516, 2017.

[8] J. C. Chan, V. Malik, W. Jia, T. Kadowaki, C. S. Yajnik, K. H. Yoon, and F. B. Hu, “Diabetes in Asia: Epidemiology, risk factors, and pathophysiology,” JAMA, vol. 301, pp. 2129–2140, 2009.

[9] N. G. Congdon, D. S. Friedman, and T. Lietman, “Important causes of visual impairment in the world today,” JAMA, vol. 290, pp. 2057–2060, 2003.

[10] S. M. Marshall and A. Flyvbjerg, “Prevention and early detection of vascular complications of diabetes,” BMJ, vol. 333, pp. 475–480, 2006.

[11] A. Hutchinson, A. McIntosh, J. Peters, C. O’keeffe, K. Khunti, R. Baker, and A. Booth, “Effectiveness of screening and monitoring tests for diabetic retinopathy—A systematic review,” Diabet. Med., vol. 17, pp. 495–506, 2000.

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

[13] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, et al., “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” JAMA, vol. 316, pp. 2402–2410, 2016.

[14] D. S. W. Ting, C. Y. L. Cheung, G. Lim, G. S. W. Tan, N. D. Quang, A. Gan, H. Hamzah, R. Garcia-Franco, I. Y. San Yeo, S. Y. Lee, et al., “Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes,” JAMA, vol. 318, pp. 2211–2223, 2017.

[15] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, “Detection of blood vessels in retinal images using two-dimensional matched filters,” IEEE Trans. Med. Imaging, vol. 8, no. 3, pp. 263–269, 1989.

[16] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, “Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images,” Br. J. Ophthalmol., vol. 83, pp. 902–910, 1999.

[17] C. E. Baudoin, B. J. Lay, and J. C. Klein, “Automatic detection of microaneurysms in diabetic fluorescein angiography,” Rev. Epidemiol. Sante Publique, vol. 32, pp. 254–261, 1984.

[18] N. Silberman, K. Ahrlich, R. Fergus, and L. Subramanian, “Case for automated detection of diabetic retinopathy,” in Proc. AAAI Spring Symp. Series, Stanford, CA, USA, Mar. 22–24, 2010. [Online]. Available: https://nyuscholars.nyu.edu/en/publications/case-for-automated-detection-of-diabetic-retinopathy

[19] D. G. Lowe, “Distinctive image features from scale-invariant key points,” Int. J. Comput. Vis., vol. 60, pp. 91–110, 2004.

[20] T. Lindeberg, “Scale-invariant feature transform,” Scholarpedia, vol. 7, p. 10491, 2012.

[21] R. Pires, H. F. Jelinek, J. Wainer, E. Valle, and A. Rocha, "Advancing bag-of-visual-words representations for lesion classification in retinal images," PLoS ONE, vol. 9, p. e96814, 2014, doi: 10.1371/journal.pone.0096814.

[22] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in European Conference on Computer Vision, Berlin, Heidelberg: Springer, 2006, pp. 404–417.

[23] L. Séoud, T. Faucon, T. Hurtut, J. Chelbi, F. Cheriet, and J. P. Langlois, "Automatic detection of microaneurysms and haemorrhages in fundus images using dynamic shape features," in Proc. IEEE 11th Int. Symp. Biomed. Imaging (ISBI), Beijing, China, Apr. 2014, pp. 101–104.

[24] L. Seoud, J. Chelbi, and F. Cheriet, "Automatic grading of diabetic retinopathy on a public database," in Ophthalmic Medical Image Analysis Int. Workshop, Iowa City, IA, USA: Univ. of Iowa, 2015.

[25] M. J. Paranjpe and M. N. Kakatkar, "Automated diabetic retinopathy severity classification using support vector machine," Int. J. Res. Sci. Adv. Technol., vol. 3, pp. 86–91, 2013.

[26] A. Liaw and M. Wiener, "Classification and regression by randomForest," R News, vol. 2, pp. 18–22, 2002.

[27] T. Joachims, Making large-scale SVM learning practical, Tech. Rep., Univ. of Dortmund, 1998. [Online]. Available: https://www.econstor.eu/handle/10419/77178

[28] T. Walter, J. C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy—detection of exudates in color fundus images of the human retina," IEEE Trans. Med. Imaging, vol. 21, pp. 1236–1243, 2002.

[29] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in European Conf. Computer Vision (ECCV), Zurich, Switzerland, Sep. 2014, Cham: Springer, pp. 818–833.

[30] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," in Proc. Workshops 32nd AAAI Conf. Artif. Intell., New Orleans, LA, USA, Feb. 2017.

[31] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in Int. Conf. Med. Image Computing and Computer-Assisted Intervention (MICCAI), Cham: Springer, 2017, pp. 533–540.

[32] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," Neurocomputing, vol. 149, pp. 708–717, 2015.

[33] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Adv. Neural Inf. Process. Syst., vol. 25, pp. 1097–1105, 2012.

[34] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in Proc. 31st AAAI Conf. Artif. Intell., San Francisco, CA, USA, Feb. 2017.

[35] O. M. Parkhi, A. Vedaldi, and A. Zisserman, Deep Face Recognition, Univ. of Oxford, 2015. [Online]. Available: https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf

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

[37] C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, Jun. 2015, pp. 1–9.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[39] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," Comput. Electr. Eng., vol. 72, pp. 274–282, 2018.

[40] 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, p. 3704, 2021.

[41] M. K. Behera, R. Mishra, A. Ransingh, and S. Chakravarty, "Prediction of different stages in diabetic retinopathy from retinal fundus images using radial basis function based SVM," Indian J. Sci. Technol., vol. 13, pp. 2030–2040, 2020. [CrossRef]

[42] M. Sankar, K. Batri, and R. Parvathi, "Earliest diabetic retinopathy classification using deep convolution neural networks," Int. J. Adv. Eng. Technol., vol. 10, p. M9, 2016.

[43] G. Lim, M. L. Lee, W. Hsu, and T. Y. Wong, "Transformed representations for convolutional neural networks in diabetic retinopathy screening," in Proc. Workshops at the 28th AAAI Conf. on Artificial Intelligence, Québec City, QC, Canada, Jul. 2014.

[44] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," Procedia Comput. Sci., vol. 90, pp. 200–205, 2016. [CrossRef]

[45] Kaggle, "Diabetic retinopathy detection challenge," 2016. [Online]. Available: http://www.kaggle.com/c/diabetic-retinopathy-detection/data. [Accessed: Jun. 1, 2022].

[46] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 2921–2929.

[47] M. Lin, Q. Chen, and S. Yan, "Network in network," arXiv preprint arXiv:1312.4400, 2013.

[48] S. A. Lekshmi and G. M. Rajathi, "Detection of glaucoma and diabetic retinopathy using image processing technique by Raspberry Pi," Indian J. Sci. Technol., vol. 12, p. 29, 2019. [CrossRef]

[49] C. H. Bindu and G. S. Sravanthi, "Retinopathy detection of eye images in diabetic patients," Indian J. Sci. Technol., vol. 12, p. 3, 2019. [CrossRef]

[50] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, J. Ma, and K. Wang, "Image preprocessing in classification and identification of diabetic eye diseases," Data Sci. Eng., vol. 6, pp. 455–471, 2021. [CrossRef]

[51] A. G. P. Henry and A. Jude, "Convolutional neural-network-based classification of retinal images with different combinations of filtering techniques," Open Comput. Sci., vol. 11, pp. 480–490, 2021. [CrossRef]

[52] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," in Proc. 2018 IEEE 15th Int. Symp. Biomedical Imaging (ISBI), Washington, DC, USA, Apr. 2018, pp. 289–293.

[53] University of Erlangen-Nuremberg, High Resolution Fundus (HRF) Image Database. Erlangen, Germany: Univ. Erlangen-Nuremberg, 2014. [Online]. Available: http://www5.cs.fau.de/research/data/fundus-images/. [Accessed: May 2, 2021].

[54] M. Goldbaum, Structured Analysis of the Retina, 2013. [Online]. Available: http://www.ces.clemson.edu/ahoover/stare/. [Accessed: May 24, 2021].

[55] DIARETDB0 Database, Medical Image Understanding and Analysis, Mar. 2014. [Online]. Available: http://www.it.lut.fi/project/imageret/diaretdb. [Accessed: Mar. 24, 2022].

[56] E. Decenciere, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, and C. Trone, "Feedback on a publicly distributed image database: The MESSIDOR database," Image Anal. Stereol., vol. 33, pp. 231–234, 2014. [CrossRef]

[57] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," Inf. Sci., vol. 501, pp. 511–522, 2019. [CrossRef]

[58] Y. P. Liu, Z. Li, C. Xu, J. Li, and R. Liang, "Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network," Artif. Intell. Med., vol. 99, p. 101694, 2019. [CrossRef]

[59] 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 Proc. 2019 41st Annu. Int. Conf. IEEE Eng. Medicine and Biology Society (EMBC), Berlin, Germany, Jul. 2019, pp. 2045–2048.

[60] S. Das, K. Kharbanda, M. Suchetha, R. Raman, and E. Dhas, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," Biomed. Signal Process. Control., vol. 68, p. 102600, 2021. [CrossRef]

[61] X. Wang, Y. Lu, Y. Wang, and W. B. Chen, "Diabetic retinopathy stage classification using convolutional neural networks," in Proc. 2018 IEEE Int. Conf. Information Reuse and Integration (IRI), Salt Lake City, UT, USA, Jul. 2018, pp. 465–471.

[62] S. H. Khan, Z. Abbas, and S. D. Rizvi, "Classification of diabetic retinopathy images based on customised CNN architecture," in Proc. 2019 Amity Int. Conf. Artificial Intelligence (AICAI), Dubai, UAE, Feb. 2019, pp. 244–248.

[63] T. Shanthi and R. S. Sabeenian, "Modified AlexNet architecture for classification of diabetic retinopathy images," Comput. Electr. Eng., vol. 76, pp. 56–64, 2019. [CrossRef]