Bulletin of Electrical Engineering and Informatics
Vol. 14, No. 4, August 2025, pp. 2752~2761
ISSN: 2302-9285, DOI: 10.11591/eei.v14i4.7854

CLAHE-AlexNet optimized deep learning model for accurate detection of diabetic retinopathy

Swetha G.[1], Gaurav Gupta[2], Kantilal Pitambar Rane[3], Omkar M. Ghag[4], Sachin K. Korde[5], Sachin Lalar[6], Batyrkhan Omarov[7], Abhishek Raghuvanshi[8]

[1]Department of Computer Science and Engineering, R R Institute of Technology, Bangalore, India
[2]G L Bajaj Institute of Technology and Management Greater Noida, Uttar Pradesh, India
[3]Department of Electronics and Telecommunications Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India
[4]Department of Artificial Intelligence and Machine Learning, Universal College of Engineering, Near Bhajanlal Dairy & Punyadham, Vasai, India
[5]Department of Computer Engineering, Pravara Rural Engineering College, Maharashtra, India
[6]Department of Engineering and Technology, Gurugram University, Gurugram, India
[7]Al-Farabi Kazakh National University, International Information Technology University, Almaty, Kazakhstan
[8]Department of Computer Science Engineering, Mahakal Institute of Technology, Ujjain, India

Article Info

Article history:

Received Sep 9, 2023
Revised Jan 8, 2025
Accepted Mar 9, 2025

Keywords:

Accuracy
AlexNet
Deep learning
F1 measure
Fundus image classification
ResNet-50
Retinopathy detection

ABSTRACT

Diabetic retinopathy (DR) is a disease that affects the blood vessels that are located in the retina. Loss of vision due to diabetes is a common consequence of the illness and a key factor in the progression of vision loss and blindness. Both ophthalmology and diabetes research have become more dependent on computer vision and image processing techniques in recent years. Fundus photography, also known as a fundus image, is a method that may be used to capture an image of the back of a person's eye. This article presents optimized deep learning model for diagnostic marking in retinal fundus images towards accurate detection of retinopathy. For experimental work, 500 images were selected from available open source Kaggle data set. 400 images were used to train deep learning model and remaining 100 images were used to validate the model. Images were enhanced using the contrast limited adaptive histogram equalization (CLAHE) algorithm. Pre trained convolutional neural network (CNN) models-AlexNet, VGG16, GoogleNet, and ResNet are used for classification and prediction of images. Accuracy, specificity, precision and F1-score of AlexNet is better than VGG16, ResNet-50, and GoogleNet. Sensitivity of ResNet-50 is higher than other pre trained CNN models.

Corresponding Author:

Abhishek Raghuvanshi
Department of Computer Science Engineering, Mahakal Institute of Technology
Ujjain, India
Email: abhishek14482@gmail.com

1. INTRODUCTION

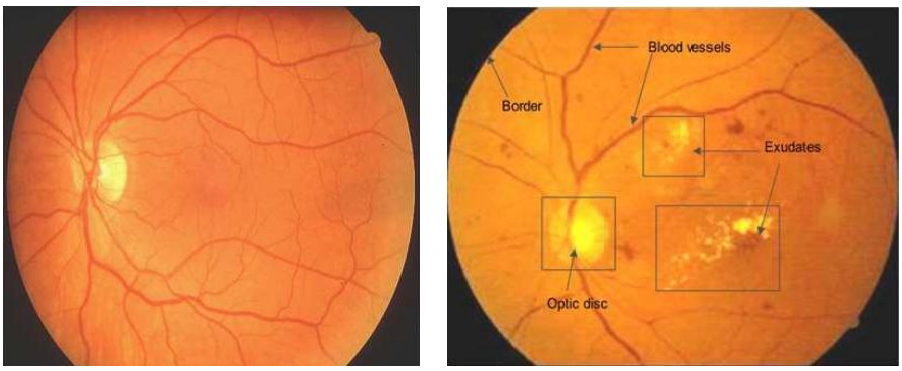
The metabolic conditions that are collectively referred to be diabetes, also often referred to as diabetes mellitus, have the same core cause, which is prolonged exposure to high blood glucose levels. Uncontrolled diabetes may lead to a number of serious complications, some of the most serious of which include damage to the kidneys, eyes, feet, heart, skin, and hearing, as well as stroke and diabetic neuropathy. Diabetic eye illness encompasses a wide range of conditions, including macular edoema, glaucoma, cataracts, and diabetic retinopathy (DR). DR is a disease that affects the blood vessels that are located in the retina, which is the light-sensitive tissue located at the back of the eye [1]. Loss of vision due to diabetes is a

common consequence of the illness and a key factor in the progression of vision loss and blindness in those under the age of 70. The most common reason for DR is a sudden spike in blood sugar, which, in turn, causes damage to the endothelium that lines blood vessels and promotes permeability in the blood vessels that supply the retina. As DR worsens, a detachment of the retina will eventually ensue [2]. This therapy choice is the least desired of the available options since persons who have DR often do not exhibit any warning signals prior to the beginning of the visual loss. In its early stages, DR may be treated using laser photocoagulation, which offers the possibility of preventing irreversible vision loss.

Both nonproliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR) are possible phases of diabetic retinopathy. The term "NPDR" describes the lack of symptoms or the minimal expression of symptoms that are typical of a disease in its early stages. Patients with diabetes often have NPDR. Diabetes causes the blood vessels in the retina to leak, which causes the retina to enlarge and results in vision loss. Vision may become blurry as a result of the accumulation of exudates in the retina, which prevents blood from reaching the macula [3]. The severity of NPDR might vary from moderate to severe depending on the individual case. Loss of central vision is the consequence of PDR, which happens when new blood vessels grow in the retina and then leak into the vitreous. Vision impairment and blindness brought on by PDR are the results of two primary causes: damage to the blood vessel and the optic nerve. PDR is a debilitating illness that may lead to complete vision loss in either eye, or perhaps both eyes. When it comes to the field of digital image processing, a still image or moving video may function in either the input or the output role of the process. The widespread use highlights the significance of digital photo processing. Both ophthalmology and diabetes research have become more dependent on computer vision and image processing techniques in recent years.

Fundus photography [4], [5], also known as a fundus image, is a method that may be used to capture an image of the back of a person's eye. It is shown in Figure 1. Figure 1(a) represents a normal image whereas Figure 1(b) represents an image having retinopathy symptoms. Poor illumination, subject movement during the exposure, longer shutter durations, and the existence of strong corneal reflexes all contributed to the blurriness of early fundus photos. The fundus camera's capabilities expanded dramatically throughout the course of its history, eventually giving rise to procedures such as non-mydriatic imaging, automated eye alignment, electrical control of illumination, and the capture of high-resolution digital pictures. As a result of recent developments in fundus photography, the documenting of retinal sickness has become an integral element of the routine ophthalmic practice. Taking a snapshot of the fundus needs a high-tech microscope that comes equipped with its own light and camera. When doing fundus photography, it is possible to see the retina, macula, and optic disc all at the same time. Analyses of fundus photography may be performed with the use of dyes such as fluorescein and indocyanine green, as well as with the application of coloured filters. Images of the fundus may be used to identify a variety of conditions, including DR, macular degeneration, glaucoma, retinal tumours, multiple sclerosis, and choroid abnormalities [6], [7].

This section presents optimized deep learning model for diagnostic marking in retinal fundus images towards accurate detection of retinopathy. For experimental work, 500 images were selected from available open source Kaggle dataset. 400 images were used to train deep learning model and remaining 100 images were used to validate the model. Images were enhanced using the contrast limited adaptive histogram equalization (CLAHE) algorithm. Pre trained convolutional neural network (CNN) models-AlexNet, VGG16, GoogleNet, and ResNet are used for classification and prediction of images. Accuracy, specificity, precision and F1-score of AlexNet is better than VGG16, ResNet-50, and GoogleNet. Sensitivity of ResNet-50 is higher than other pre trained CNN models.



(a) (b)

Figure 1. Fundus images; (a) normal image and (b) retinopathy image

2. LITERATURE SURVEY

In this study, authors have explained how CNN may be used to accurately differentiate between signals in optical coherence tomography that are suggestive of PDR and those that are not indicative of PDR [8]. During the preprocessing stage, patches of retinal pictures are removed from the overall picture so that they may be used to train a CNN. These extracted patches are then used in the subsequent step. In order to accomplish maximum productivity, we make use of managers of transfer studies and the integration of various aspects. In addition to that, the planned implementation of the system is broken out in great depth. The results for a pre-trained CNN system that used nasal and temporal retina patches independently illustrate how this may be done via the evaluation of several alternative system set-up conditions and the selection of the optimal one.

It is noted in photographic fundus that the capacity to ameliorate difficulties caused by gravity and DR is present [9]. In order to establish the existence of connected components and define the severity of DR grades, the first thing we do is study the consistency of grading by looking at the variability of the various graders. This helps us determine the severity of DR grades. In addition, we take an impartial approach to comparing and contrasting the various annotation techniques to severity level predictions. The recommended annotations of data are shown here, along with the severity levels and whether or not linked characteristics are present.

Research by Costa *et al.* [10] demonstrate a system that can detect DR from retinal pictures with just little human oversight. By simultaneously optimizing the encoding instance and the image classification phase, the author achieves a considerable improvement over previous methods that are presently in use. The acquisition of valuable intermediate-level pathological images is made possible as a result of this. It is necessary for the loss function that is used in this technique to have correct instance and mid-level presentations in order to further improve the explain ability of the model's judgements.

There are certain guidelines offered in for the use of the platform function bag in the detection of DR gravitational rates [11]. A newly developed and improved automated method for determining the severity of DR, this method makes use of a dictionary-based approach and does not need any pre- or post-processing measures. This approach is dependent on providing an accurate portrayal of the pathological picture within the context of a classroom environment. Quantifying the descriptive features of retinal visuals is accomplished via the use of algorithms and histograms of directed gradients. These results are then included into the construction of a visual function dictionary. Features such as these are compiled into a dictionary, where they are then subjected to coding and pooling in order to provide a more condensed feature representation.

Due to the severity and close geographical proximity of DR and malarial retinopathy, it has become increasingly crucial to detect leaks. The image is initially super pixelized in order to construct a hierarchy of patches that are important to the overall effect. In order to combine the saliency maps that are comparable at the same stage of development, an average operator that takes into account the salience indices of strength and compactness is used. Siamese-like CNN automation detecting DR is something that is covered in [12]. Transfer learning is a strategy that has emerged as one of the most cutting-edge advancements in the area of CNN in recent years. The model is an improvement over previous studies in that it is able to utilize photographs of the binocular fundus as inputs, and it also knows how to cater for them.

Research by Seoud *et al.* [13] outline how to employ dynamic shape features for the purpose of identifying red lesions in the context of DR screening. The most important addition is a new set of form characteristics that have been given the name Dynamic Shape Characteristics. These characteristics do not need the precise segmentation of any areas. These properties allow for the differentiation of vessel segments and lesions, as well as the provision of shape data during photo floods. In this article, a short discussion is provided on pre-processing and diagnostic methods for the detection of DR in human eyes. The pre-processing phase consists of a variety of steps, such as photo rectification, green channel extraction, image contrast enhancement, noise reduction filtering, Gabor filtration, vessel calculation, pressure test, and other moments of invariance-related steps. Artificial neural network (ANN) classifier is used to complete the classification.

There are many processes involved in DR detection, including image processing, segmentation, extraction, and classification. The detection of DR with the use of a deep learning classification system is described in [14]. Processing the images captured by the retina must come first in order to ensure that the results are reliable. The evaluation of the image's overall quality is the topic of discussion in this part of the course. After then, a large number of photographs with enlarged dimensions are generated at random. During the final step, the findings from a variety of tests are aggregated into a single complete evaluation. When it is possible to do so, adding the total ratings from all of the other eyes results in greater accuracy.

Research by Rakhlin [15] describes the development and validation of a deep learning system for DR detection. The deep learning computer family does away with the need to explicitly describe rules by

having an algorithm learn from a large number of examples that mimic the behaviour that is supposed to be produced by the programmed. The use of colour fundus pictures for the diagnosis of DR and the detection of morphological exudates. The photo goes via a greyscale converter to be processed. The preliminary processing continues with the normalization of the histogram as one of the steps. The thresholding approach is used for the purpose of accomplishing this goal by converting continuous pictures into discrete binary ones. It is possible for there to be morphological alterations, such as an increase in size or a decrease in size as a result of erosion or dilatation. Therefore, a distance measurement and a modification of the watershed are required for the identification of exudates.

Kaur *et al.* [16] provide a computer algorithm in that may be used to detect PDR in fundus colour photographs. The first thing that needs to be done is to categories the patches according to their respective textures. The properties that are determined by the vessels themselves are eliminated. Last but not least, a rule-based classifier is put to use so that the actual classifications may be made. An ANN may be used to identify DR in any of its various stages, as described. The image that was provided is converted to grayscale as the first stage. After that, we use digital filters that segment based on the magnitude of the gradient in order to accomplish histogram equalization via fuzzy clustering. The mean and the total of the pixels are both used in the process of feature extraction. In conclusion, an ANN in conjunction with a multilayer perceptron is used for classification.

3. METHOD

This section presents optimized deep learning model for diagnostic marking in retinal fundus images towards accurate detection of retinopathy. For experimental work, 500 images were selected from available open source data set [17]. 400 images were used to train deep learning model and remaining 100 images were used to validate the model. Images were enhanced using the CLAHE algorithm. Pre trained CNN models-AlexNet, VGG16, GoogleNet, and ResNet are used for classification and prediction of images.

The retinal photographs that were obtained during the exams suffer from a number of issues, including poor contrast and inconsistent lighting, to name a few of these issues. Images of the fundus retina seem to have a higher degree of brightness closer to the centre of the retina compared to its more peripheral regions. The regions that make up the image's edges and borders are not lighted at all, in contrast to the sections that make up the middle of the picture, which is well illuminated. As a consequence of this, some lesions may not be identified because there is not enough contrast amplification in the picture of the retina. This is because the image's centre as well as its overall density (OD) have both been enhanced digitally. CLAHE is an acronym that stands for "contrast limited adaptive histogram equalization," and it refers to a technique that is used to increase the contrast of a picture throughout the whole retina. Using a technique known as contrast limited adaptive histogram equalization, or CLAHE for short, the histogram is split into two sections with dimensions that are comparable to one another [18]. CLAHE, which stands for contrast limited adaptive histogram equalization, is a technique for improving the contrast of a picture. To begin, the section of the image that needs its contrast improved is segmented into multiple smaller pieces of the same size. After then, the contrast in every one of these areas is modified in its own unique way. The process of applying histogram equalization to each region alters the way that greyscale value is distributed over the region, so exposing features that were previously hidden. Utilizing this technique results in an overall improvement to the picture's quality. If CLAHE is added to the green channel of a fundus image, it may make blood vessels, microaneurysms, hemorrhages, and exudates more visible [19].

The VGG object recognition model is considered to be state-of-the-art and is capable of supporting up to 19 layers. In addition to ImageNet, the deep CNN architecture that VGG uses provides higher outcomes than the baseline in a wide variety of applications and datasets. VGG continues to be one of the most widely used image-recognition frameworks that are currently accessible. The number 16 denotes the weight of each layer in the VGG16 structure, which has a total of sixteen layers. This enormous network may be described using 138,000,000 different attributes in total. Even by the standards of today, it is a substantial sum of money. On the other hand, the VGG16 architecture stood out due to its simple structure, which made it highly appealing. The design seems to comply to a very strict level of consistency, as seen by the observer. Following the application of many convolutional layers, the volume is then compressed using a pooling layer in order to make it a more manageable size [20]. The pre-trained vgg16 architecture has a total of 41 layers in its construction. This design calls for an input layer that is 224 squares on each side and three rows deep, with a stride of 1 and padding of 1. The proposed network has a total of 16 convolutional layers, each of which has a kernel size of 3*3 and a stride of 1, as well as 5 Maxpooling layers, each of which has a kernel size of 2*2 and a stride of 1. In the end, there are three layers that are linked to one another, and each of these levels has two nodes with 4,096 channels and one node with 1,000 channels that may be altered as necessary dependent on the number of categorization layers (2 or 5). VGG 16 is shown in Figure 2 [21].

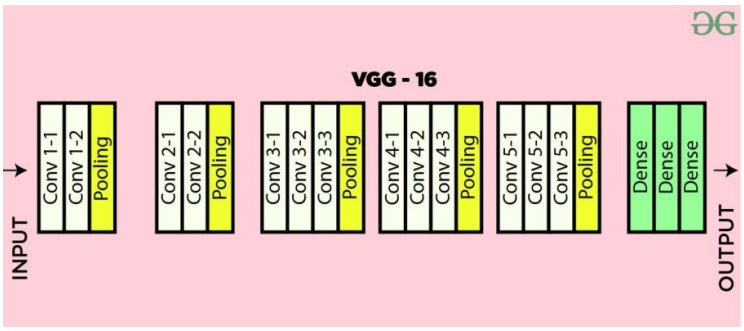


Figure 2. VGG 16 pre trained CNN

In the discipline of machine learning, particularly in the arena of deep learning's application to machine vision, the CNN that goes by the name AlexNet has had far-reaching repercussions. AlexNet is a very powerful model that can achieve outstanding performance even when used to the most difficult datasets. On the other hand, the performance of AlexNet suffers a significant setback if one of its convolutional layers is eliminated [22]. AlexNet, a leading architecture for any object-detection activity, may have applications in computer vision and artificial intelligence difficulties. These challenges can range from simple to complex. In the field of image processing, CNNs could be obsolete by AlexNet in the near future. The development of AlexNet was a significant step in expanding the applicability of deep learning to other fields, such as natural language processing and medical image analysis. AlexNet is prepared to begin its job immediately because to the 25-layer architecture it utilizes. The network has a 227 by 227 by 3 input layer that has been zeroed out for its centre normalization. This structure makes use of a cross-channel normalization that has five channels for each element, two 11*11 convolutional layers that have stride 2 and zero padding, and four 3*3 grouped convolutions that have stride 1 and padding. The pooling layer of the network is implemented using Maxpooling and has 33 kernels, 22 stride, and 00 padding. Last but not least, the classification layers (either 2 or 5) decide which of the three fully connected layers is used, with each layer consisting of two 4,096 channels and one 1,000 channels [23].

An acronym for "Residual Network," which is often referred to by its shorter form, "ResNet." There are several distinct iterations of ResNet [24], each of which is based on the same notion but employs a different number of layers. All of these iterations of ResNet may be found online. A neural network may have up to 50 layers if it is implemented using the Resnet-50 variant of the algorithm. The need to find a solution to this issue was one of the driving forces for the development of ResNet. Architecture of ResNet-50 is shown in Figure 3. Deep residual nets make use of residual blocks in order to increase the accuracy of the models they create. This neural network's value rests in its residual blocks, and more precisely in the concept of "skip connections." You have two options available to you for making advantage of these skip connections. They begin the process of resolving the issue of vanishing slopes by building a new path for the gradient to follow. The combination of these characteristics enables the model to gain the capability of learning an identity function [25]. These characteristics are as follows: Because of this, the functionality of the model's top layers will always be assured to be at least as good as that of its lower ones. To summarize, the residual blocks make it easier for the layers to learn the identity functions they need to perform. Because of this, ResNet is able to construct larger neural networks with more neural layers more efficient while also reducing the percentage of errors. In other words, the skip connections allow for even deeper networks to be trained by combining the outputs of previous layers with the outputs of stacked levels. This is done by combining the outputs of the stacked levels with the outputs of the prior layers. The pretrained Resnet-50 network is comprised of a total of 50 layers. A 224 by 224 input is required for Resnet-50. Each convolutional block is made up of three convolutional layers that are 7×7 in size, and one Maxpooling layer that is 3×3. Not least of all, there are three layers that are entirely connected, each having two 4,096-channel and one 1,000-channel sublayers that may be modified as required depending on the classification sublayers (2 or 5).

Image recognition and classification are handled using the GoogleNet [22] architecture, which is a deep learning CNN architecture. A CNN architecture, similar to neural network design, is built up of neurons, the weights and biases of which may be taught. Every neuron takes in information from a variety of different places, computes a weighted sum, employs an activation function, and then sends the result out into the world. The same loss function is applied to each node that makes up a CNN, and the techniques that were established for neural networks may still be used with CNNs. In GoogleNet, an inception module is used as the base, and succeeding layers are created on top of one another. In order to improve GoogleNet's computational efficiency, each layer applies parallel filter operations to the input of the layer below it. These designs are often

implemented using graphics processing units, also known as GPUs. GPUs are developed expressly to speed up labor-intensive computations such as image weight multiplication, and thus are commonly used to execute these designs. The difficulty with these designs is that each computation takes a significant number of memory accesses. This not only results in an increase in the amount of power that is used, but also makes it problematic to employ these designs in mobile apps that do not have specialized accelerator hardware.

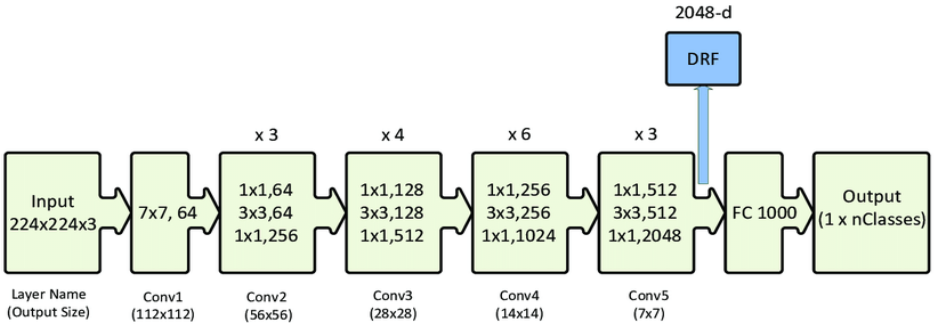


Figure 3. Architecture of ResNet-50

4. RESULTS AND DISCUSSION

For experimental work, 500 images were selected from available open source data set [17]. 400 images were used to train deep learning model and remaining 100 images were used to validate the model. Images were enhanced using the CLAHE algorithm. Pre trained CNN models- AlexNet, VGG16, GoogleNet, and ResNet are used for classification and prediction of images. The results are shown in Table 1 and Figure 4. Accuracy, specificity, precision and F1-score of AlexNet is better than VGG16, ResNet-50, and GoogleNet. Sensitivity of ResNet-50 is higher than other tre prained CNN models.

Table 1. Confusion matrix

Parameter	GoogleNet	ResNet	VGG16	AlexNet
TP	210	225	227	235
TN	127	143	144	153
FP	38	28	12	4
FN	28	4	17	8

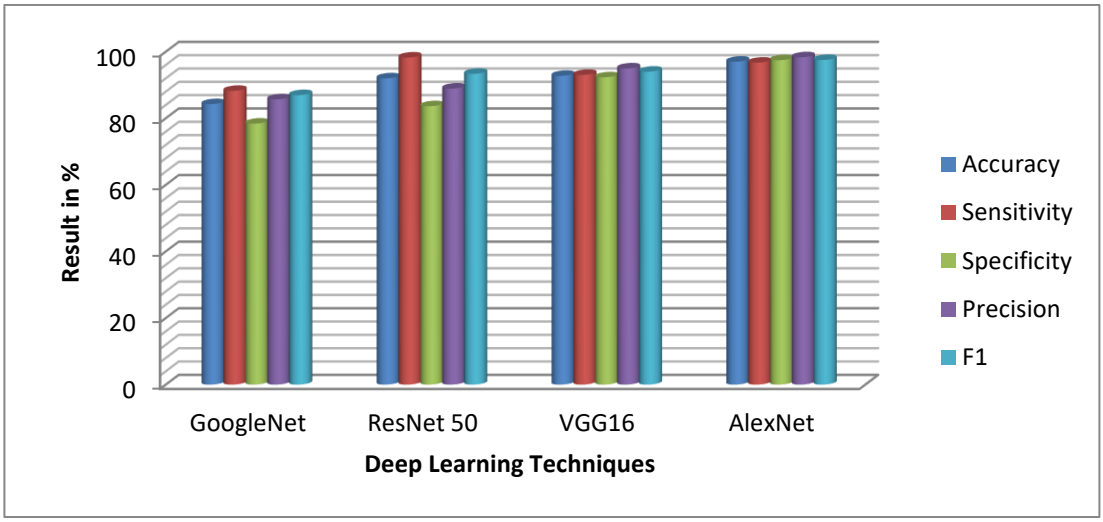


Figure 4. Performance comparison for DR classification and prediction

In the suggested model, five metrics such as accuracy, specificity, sensitivity, precision, and F1-score are utilized to assess the classification performance of the classification phase. Accuracy is the ratio of the number of correctly labeled samples to the whole pool of samples. It is defined as,

$$accuracy = (tp + tn)/N$$

Specificity is the measure used to assess a model's ability to predict true negatives in each given category. It is the ratio of the number of true negative evaluations to the total number of negative evaluations.

$$specificity = tn\_rate$$

Sensitivity is the measure used to assess a model's ability to predict true positives in each given category. It is the ratio of the number of true positive evaluations to the total number of positive evaluations.

$$sensitivity = tp\_rate$$

Precision represents the proportion of true positives to all positives.

$$precision = tp/(tp + fp)$$

F1 measure represents the average of precision and recall. It lends more weight to false positives and false negatives, and it is an overall measure of a model's correctness;

$$f\_measure = 2 * ((precision * recall)/(precision + recall))$$

where; true positive (TP) (if a disease is shown to be existing in a patient and the provided diagnostic test also indicates the disease presence, the diagnostic test outcome is referred to as true positive); true negative (TN) (if a disease is demonstrated to be non-existent in a patient and the diagnosis test likewise confirms that the disease is non-existent, the test outcome is referred to as true negative); false positive (FP) (if a diagnostic test detects disease in a patient who does not have the disease, the test outcome is termed false positive); and false negative (FN) (if the findings of a diagnostic test show that the disease is not present in a patient who has the condition, the test outcome is referred to as false negative).

5. CONCLUSION

Diabetes may cause vision impairment, which is a frequent complication of the disease and a significant contributor to the decline in visual acuity that can lead to blindness in those younger than 70 years old. The most typical cause of DR is a fast rise in blood sugar, which, in turn, causes damage to the endothelium that lining blood vessels and increases permeability in the blood vessels that feed the retina. This is the most prevalent cause of the condition. The detachment of the retina will ultimately occur if the DR is not treated in time. This article provides a deep learning model that has been tuned for diagnostic marking in retinal fundus pictures, with the goal of more accurately detecting retinopathy. In the course of the experimental effort, 500 photos were chosen from the open source Kaggle data set that was readily accessible. In order to train the deep learning model, 400 photographs were employed, while the remaining 100 images were used to verify the model. The CLAHE algorithm was used to improve the quality of the images. For the purposes of picture classification and prediction, the CNN models AlexNet, VGG16, GoogleNet, and ResNet that have already undergone training are used. AlexNet outperforms VGG16, ResNet-50, and GoogleNet in terms of accuracy, specificity, and precision, and it also has a higher F1-score. The ResNet-50 model's sensitivity is much greater than that of previous trained CNNs.

ACKNOWLEDGMENTS

We are deeply thankful to our institutions for providing the necessary resources and facilities that enabled the successful completion of this research.

FUNDING INFORMATION

No funding received for this research work.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Swetha G.	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Gaurav Gupta		✓				✓		✓	✓	✓	✓	✓		
Kantilal Pitambar Rane	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	
Omkar M. Ghag	✓	✓		✓		✓		✓	✓	✓			✓	
Sachin K. Korde		✓				✓		✓	✓	✓	✓	✓		
Sachin Lalar	✓	✓	✓	✓		✓	✓			✓	✓		✓	✓
Batyrkhan Omarov	✓		✓	✓	✓	✓		✓	✓	✓			✓	
Abhishek Raghuvanshi	✓		✓	✓	✓	✓	✓		✓		✓		✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

Ethical approval was not required for this study, as it did not involve human or animal subjects.

DATA AVAILABILITY

The data that support the findings of this study are openly available in [Kaggle, "Diabetic Retinopathy 2015 Data Colored Resized," www.kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-2015-data-colored-resized] at reference number [26].

REFERENCES

[1] M. T. Esfahani, M. Ghaderi, and R. Kafiyeh, "Classification of diabetic and normal fundus images using new deep learning method," *Leonardo Electronic Journal of Practices and Technologies*, no. 32, pp. 233–248, 2018.

[2] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

[3] M. Shaban *et al.*, "A convolutional neural network for the screening and staging of diabetic retinopathy," *PLoS ONE*, vol. 15, no. 6 June, 2020, doi: 10.1371/journal.pone.0233514.

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

[5] R. K. Veluri *et al.*, "Learning analytics using deep learning techniques for efficiently managing educational institutes," *Materials Today: Proceedings*, vol. 51, pp. 2317–2320, 2022, doi: 10.1016/j.matpr.2021.11.416.

[6] V. D. P. Jasti *et al.*, "Computational Technique Based on Machine Learning and Image Processing for Medical Image Analysis of Breast Cancer Diagnosis," *Security and Communication Networks*, pp. 1–7, 2022, doi: 10.1155/2022/1918379.

[7] T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," *Future Healthcare Journal*, vol. 6, no. 2, pp. 94–98, 2019, doi: 10.7861/futurehosp.6-2-94.

[8] M. Ghazal, S. S. Ali, A. H. Mahmoud, A. M. Shalaby, and A. El-Baz, "Accurate Detection of Non-Proliferative Diabetic Retinopathy in Optical Coherence Tomography Images Using Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 34387–34397, 2020, doi: 10.1109/ACCESS.2020.2974158.

[9] J. Wang, Y. Bai, and B. Xia, "Feasibility of Diagnosing Both Severity and Features of Diabetic Retinopathy in Fundus Photography," *IEEE Access*, vol. 7, pp. 102589–102597, 2019, doi: 10.1109/ACCESS.2019.2930941.

[10] P. Costa, A. Galdran, A. Smailagic, and A. Campilho, "A Weakly-Supervised Framework for Interpretable Diabetic Retinopathy Detection on Retinal Images," *IEEE Access*, vol. 6, pp. 18747–18758, 2018, doi: 10.1109/ACCESS.2018.2816003.

[11] M. Leeza and H. Farooq, "Detection of severity level of diabetic retinopathy using Bag of features model," *IET Computer Vision*, vol. 13, no. 5, pp. 523–530, 2019, doi: 10.1049/iet-cvi.2018.5263.

[12] X. Zeng, H. Chen, Y. Luo, and W. Ye, "Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network," *IEEE Access*, vol. 7, pp. 30744–30753, 2019, doi: 10.1109/ACCESS.2019.2903171.

[13] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. M. P. Langlois, "Red Lesion Detection Using Dynamic Shape Features for Diabetic Retinopathy Screening," *IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1116–1126, 2016, doi:

10.1109/TMI.2015.2509785.
[14] P. Kaur, S. Chatterjee, and D. Singh, "Neural network technique for diabetic retinopathy detection," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6, pp. 440–445, 2019, doi: 10.35940/ijeat.E7835.088619.
[15] A. Rakhlin, "Diabetic Retinopathy detection through integration of Deep Learning classification framework," *bioRxiv*, pp. 1–11, 2018.
[16] N. Kaur, S. Chatterjee, M. Acharyya, J. Kaur, N. Kapoor, and S. Gupta, "A supervised approach for automated detection of hemorrhages in retinal fundus images," in *2016 5th International Conference on Wireless Networks and Embedded Systems, WECON 2016*, 2017, pp. 1–5, doi: 10.1109/WECON.2016.7993461.
[17] kaggle, "Diabetic Retinopathy 2015 Data Colored Resized," www.kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-2015-data-colored-resized. (Date accessed: Sept. 10, 2023).
[18] M. Ravikumar, P. G. Rachana, B. J. Shivaprasad, and D. S. Guru, "Enhancement of Mammogram Images Using CLAHE and Bilateral Filter Approaches," *Cybernetics, Cognition and Machine Learning Applications: Proceedings of ICCCMLA 2020. Springer Singapore*, pp. 261–271, 2021, doi: 10.1007/978-981-33-6691-6_29.
[19] M. Hayati *et al.*, "Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning," *Procedia Computer Science*, vol. 216, pp. 57–66, 2022, doi: 10.1016/j.procs.2022.12.111.
[20] C. Sitaula and M. B. Hossain, "Attention-based VGG-16 model for COVID-19 chest X-ray image classification," *Applied Intelligence*, vol. 51, no. 5, pp. 2850–2863, 2021, doi: 10.1007/s10489-020-02055-x.
[21] N. Javaid, "A PLSTM, AlexNet and ESNN based ensemble learning model for detecting electricity theft in smart grids," *IEEE Access*, vol. 9, pp. 162935–162950, 2021, doi: 10.1109/ACCESS.2021.3134754.
[22] M. Ronald, A. Poulose, and D. S. Han, "ISPLInception: An Inception-ResNet Deep Learning Architecture for Human Activity Recognition," *IEEE Access*, vol. 9, pp. 68985–69001, 2021, doi: 10.1109/ACCESS.2021.3078184.
[23] G. H. Babu and N. Venkatram, "An efficient image dahazing using Googlenet based convolution neural networks," *Multimedia Tools and Applications*, vol. 81, no. 30, pp. 43897–43917, Dec. 2022, doi: 10.1007/s11042-022-13222-2.
[24] H. C. Chen *et al.*, "AlexNet Convolutional Neural Network for Disease Detection and Classification of Tomato Leaf," *Electronics*, vol. 11, no. 6, pp. 1–17, Mar. 2022, doi: 10.3390/electronics11060951.
[25] Y. R., V. R. S. M., R. Panjanathan, G. J. S, and J. A. L, "Diabetic Retinopathy Classification Using CNN and Hybrid Deep Convolutional Neural Networks," *Symmetry*, vol. 14, no. 9, pp. 1–14, Sep. 2022, doi: 10.3390/sym14091932.

BIOGRAPHIES OF AUTHORS

Dr. Swetha G. is working as Associate Professor in the Department of Computer Science and Engineering in R R Institute of Technology in Bangalore, India. She is having a teaching experience of 15 years. Her research interests are deep learning and machine learning. She can be contacted at email: swetha.ganganala@gmail.com.

Gaurav Gupta is an Assistant Professor from 2013 in the Department of Computer Science and Engineering in G.L. Bajaj Institute of Technology & Management, Greater Noida. He has done his MCA from Ajay Kumar Garg Engineering College, Ghaziabad M.Tech. from Monad University, Hapur. He can be contacted at email: gaurav.mintu@gmail.com.

Dr. Kantilal Pitambar Rane received B.E. degree in Electronics and Telecommunication Engineering from North Maharashtra University, Jalgaon in 1996. M.E. Degree in Electronics Engineering from Shivaji University, Kolhapur in 2005. Ph.D. Degree in Electronics Engineering from North Maharashtra University, Jalgaon in 2013. He is currently working as Professor at Bharati Vidyapeeth College of Engineering Navi Mumbai, India. His research interests include image processing, IoT, AI, and machine learning. He can be contacted at email: kantiprane@rediffmail.com.

Mr. Omkar M. Ghag is presently working as an Associate Professor in the Department of Information Technology at Sinhgad Academy of Engineering, Pune, India. He has been in teaching profession for more than 5 years. His main area of interest includes machine learning, deep learning, data science, blockchain, processor architecture, cloud computing, natural language processing, and social media analytics. He can be contacted at email: omkarghag108@gmail.com.

Dr. Sachin K. Korde is Assistant Professor in Department of Computer Engineering in Pravara Rural Engineering College in Loni, India. He can be contacted at email: kordesk99@gmail.com.

Dr. Sachin Lalar is working as Assistant Professor in Kurukshetra University in Department of Computer Science and Applications. He has received his Ph.D. in Computer Science and Engineering. He has more than 15 years' experience in teaching and research. His research interests are computer networks, data structure, and programming. He has visited many universities for presenting the research papers. Several research papers of his work published in the leading National and International Journals. He can be contacted at email: sachin509@kuk.ac.in.

Batyrkhan Omarov received his bachelor's and master's degrees from Al-Farabi Kazakh National University, Almaty, Kazakhstan in 2008 and 2010, respectively. In 2019, he received his Ph.D. from Tenaga National University, Kuala Lumpur, Malaysia. His research interests include machine learning, deep learning, data science, image processing, medical image analysis, and artificial intelligence in medicine. He can be contacted at email: batyahan@gmail.com.

Dr. Abhishek Raghuvanshi is working as head of Department in Department of Computer Science and Engineering in Mahakal Institute of Technology in Ujjain in India. He is having teaching experience of 17 years, research experience of 13 years and administrative experience of 7 years. His research areas include- machine learning, internet of things security, and health care analytics. He is having many publications in SCI and Scopus indexed journals. He has also worked on many governments of India funded research projects. He can be contacted at email: abhishek14482@gmail.com.