



# Classification of diabetic retinopathy severity level using deep learning

Santhi Durairaj<sup>1</sup> · Parvathi Subramanian<sup>2</sup> · Carmel Sobia Micheal Swamy<sup>3</sup>

Received: 20 February 2024 / Accepted: 7 March 2024 / Published online: 28 March 2024  
© The Author(s), under exclusive licence to Research Society for Study of Diabetes in India 2024

## Abstract

**Background** Diabetic retinopathy (DR) is an eye disease developed due to long-term diabetes mellitus, which affects retinal damage. The treatment at the right time supports people in retaining vision, and the early detection of DR is the only solution to prevent blindness.

**Objective** The development of DR shows few symptoms in the early stage of progression; it is difficult to identify the disease to give treatment from the beginning. Manual diagnosis of DR on fundus images is time-consuming, costly, and liable to be misdiagnosed when compared to computer-aided diagnosis systems.

**Methods** In this work, we proposed a deep convolutional neural network for the recognition and classification of diabetic retinopathy lesions to identify the severity of the disease. The performance evaluation of the proposed model was tested with other machine learning classifiers such as K-nearest neighbor (KNN), Naïve Bayes (NB), logistic regression (LR), support vector machine (SVM), decision tree (DT), and random forest (RF).

**Results** Our proposed model achieves 98.5% accuracy for the recognition and classification of the severity level of DR stages such as no DR, mild DR, moderate DR, severe DR, and proliferative DR.

**Conclusion** The training and testing of our model are carried out on images from the Kaggle APTOS dataset, and this work can act as a base for the autonomous screening of DR.

**Keywords** Diabetic retinopathy · DCNN · Machine learning classifiers · Kaggle APTOS

## Introduction

Diabetic retinopathy is a hostile impact of diabetes, which leads to vision-threatening problems if the treatment is not given at the right time. In healthcare, early detection of the disease is most important to provide better treatment. DR is a complexity of diabetes, which is the common cause of

damage to the blood vessels that give nutrition to the retina. DR causes the blood vessels to become larger and seep fluids and blood from them. DR steers to permanent vision loss if left untreated in an early stage. Around the globe, due to DR, 2.6% of blindness occurs in people who have suffered from this illness for a long time.

To decrease the rate of blindness and spread awareness about DR, Rajiv Raman et al. [1] implemented a multi-center cross-sectional screening study for people aged over 40 and more in India. The authors did the screening in ten Indian states and one union territory for the duration of 16 months from December 2018 to March 2020. During their research, 42,146 people were screened. Out of those, 18.8% of the people were diagnosed with diabetes. Their study showed that the prevalence of DR is 12.5%, and for visually threatening DR is 4.0%. According to their research, in India, 3 million people with diabetes who are aged 40 years or older live with the risk of vision loss. World Health Organization (WHO) indicated that there are 77 million people in India over the age of 18 years who are suffering from type 2 diabetes, and

✉ Santhi Durairaj  
santhijeyan@mepcoeng.ac.in

Parvathi Subramanian  
parvathi@uitkovilpatti.ac.in

Carmel Sobia Micheal Swamy  
carmelsobia@psr.edu.in

<sup>1</sup> Bio Medical Engineering, Mepco Schlenk Engineering College, Sivakasi, India

<sup>2</sup> Electronics and Communication Engineering, Unnamalai Institute of Technology, Kovilpatti, India

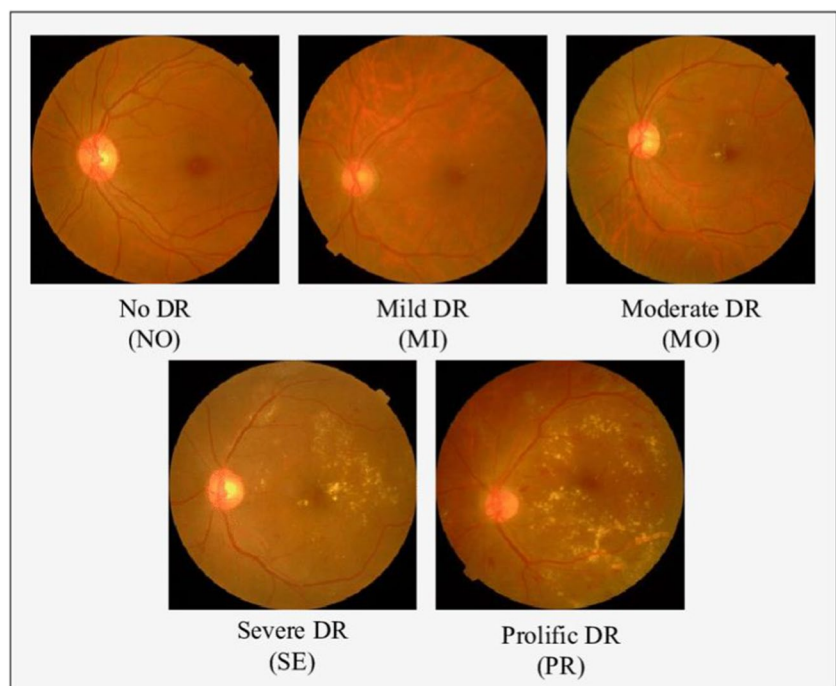
<sup>3</sup> Electrical and Electronics Engineering, PSR Engineering College, Sivakasi, India

25 million people are prediabetics, i.e., at high risk of growing diabetes population soon. More than 50% of people need to gain awareness about their disease status. Continuous retinal monitoring is most important for patients under treatment for diabetes in its early stage to prevent blindness. The disease’s severity level can be identified by detecting small lesions formed in the retina of the DR-affected patients.

The presence of these lesions, such as hard exudates (HE), cotton wool spots (or) soft exudates (CWS), microaneurysms (MA), and hemorrhages (HM), indicates the severity level of the DR. MA is the prior indication of DR with tiny red circular dots in the retina as the vessel wall weakens. The size of the MA ranges from 125 μm with

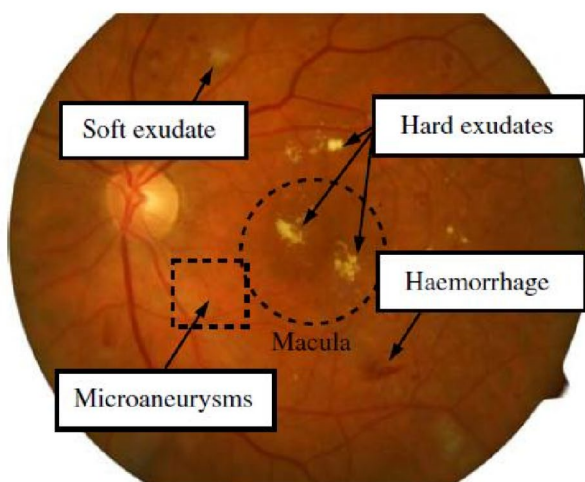
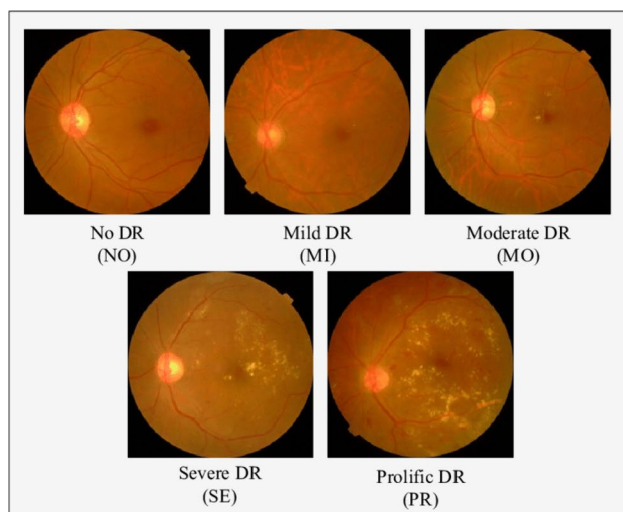
sharp margins. HM shows up with more prominent spots with a size exceeding 125 μm and irregular margins. HE presents as bright-yellow spots in the retinal outer layer, which are caused by leakage of plasma and have sharp margins. CWS seems to be white spots originating from nerve fiber swelling with round or oval shapes. Based on their appearance in color, these lesions are categorized as red (MA and HM) and bright lesions (HE and CWS). Based on the presence of these lesions, the severity level of the disease can be identified as no DR, mild DR, moderate DR, severe DR, and proliferative DR. The severity level of DR and the formation of lesions in the retina are presented in Fig. 1, Table 1, and Fig. 2, respectively.

**Fig. 1** Severity level of DR. Sample image from APTOS dataset



**Table 1** Severity level of DR based on the presence of lesions

Severity stages	DR severity level	Lesions presence
Stage 0	No DR	Absence of lesions
Stage 1	Mild non-proliferative DR	MA alone
Stage 2	Moderate non-proliferative DR	More than just MA but less than severe DR
Stage 3	Severe non-proliferative DR	Anyone: <ul style="list-style-type: none"> <li>• more than 20 intraretinal HM in each of the four quadrants</li> <li>• definite venous beading in 2+ quadrants</li> <li>• Prominent intra-retinal micro-vascular abnormalities in 1+ quadrant</li> <li>• no signs of proliferative DR</li> </ul>
Stage 4	Proliferative DR	One or more: <ul style="list-style-type: none"> <li>• Vitreous/pre-retinal HM, neovascularization</li> </ul>



**Fig. 2** Lesions formed in the retina due to DR

In the severity level, stage 1 is the earlier one, where the formation of MA can occur. In the next stage, with the advancement of the disease, the vision got blurred due to swelling in the blood vessels. In the next stage, aberrant blood vessel growth can be identified. In the severe DR stage, many blood vessels are blocked. The proliferative DR is an advanced stage where retinal detachment occurs with a retina break, leading to complete vision loss. Regular retinal screening is necessary to treat this disease in the earliest stage to retain vision.

Currently, manual diagnosis is performed with a comprehensive dilated eye exam. In this screening test, eye drops are applied to the eye to widen the pupils and get a better view of the inner part of the eyes. After that, a dye is injected into the body to check blood vessels in the eyes. By this, manual grading can be done to identify the presence or absence of DR. This manual diagnosis is time-consuming,

costly, and requires human expertise. The diagnosis of DR is easy and affordable when it is automated. In this work, the diagnosis of DR severity can be identified with deep learning techniques to make the process more efficient by saving cost and time.

The flow of this paper is organized as follows. “**Related works**” briefly discusses the related works, while “**Methodology**” presents the methodology to classify the severity level of DR using deep learning techniques. “**Experimental results**” describes the performance measure of the proposed work, while “**Performance evaluation analyses**” shows the results and discussion. The conclusion summary is presented in “**Conclusion and future work.**”

## Related works

Using machine learning models is very helpful in solving many issues related to image classification [2] in healthcare work and particularly in disease predictions [3]. A greater number of works have been proposed to detect DR using machine learning techniques [4] such as SVM [5] and decision tree [6].

Previously, various methods have been used to classify normal and abnormal retinal images with DR. Deep learning is an essential and efficient platform for performing automated and reliable decision-making tasks [7]. Deep learning models such as convolutional neural networks have been proposed for deep feature extraction and image classification applications [8–10]. Qiao et al. [11] presented their research to detect MA for the DR grading system, including mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, and severe NPDR using DCNN. Saranya and Uma Maheswari [12] proposed their work with convolution neural network (CNN) to classify images into four phases of DR.

Zhang et al. [13] and Hemanth et al. [14] proposed their work on classifying DR stages in the retinal images into four classes. However, the severity levels used in both works are different. Zhang et al. used four severity levels of DR N:o DR, mild NPDR, moderate NPDR, severe NPDR, and PDR. At the same time, Hemanth et al. presented their classification of the retinal images into normal, macular edema, PDR, and NPDR.

Based on the literature study, it is visible that developing an automated diagnosis system for detecting DR is a must to identify the disease severity at an early stage. Considering the importance, the objective of this research work is proposed. We aim to develop a robust and efficient model for the recognition and classification of DR severity stages using the Kaggle APTOS 2019 dataset [15]. The proposed work could minimize the time for effectively diagnosing the disease with less cost.

## Methods and Materials

The early detection of DR severity stages plays a vital role in better treatment. As an initial step, the dataset selection can be done. For our research work, the APTOS 2019 Kaggle dataset was used to recognize and classify the severity level of DR. The dataset images were taken using fundus photography under varying illumination conditions. These images are manually graded as class 0 to class 4 to indicate various severity levels. The number of images available in the dataset for each severity level is listed in Table 2.

From the total of 3,662 images, 75% of the images were used for training purposes, and the remaining 25% of the retinal images were used for testing purposes. Furthermore, the image preprocessing procedure includes adding more images using augmentation techniques [16] to reduce overfitting errors and noise during the training phase. This work used the following image augmentation techniques to classify DR stages. The augmentation takes place with image rotation, color transformation, width shifting and height shifting, zooming, horizontal flipping, and blurring operations to generalize the model well for recognition of DR severity stages. The database is then expanded with the augmented images and the images from the expanded database for the training and testing phase.

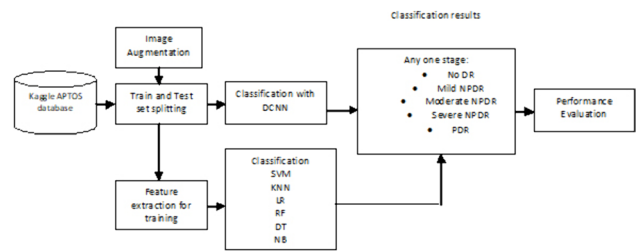
In this proposed work, the detection of the severity level of diabetes with DCNN is performed on the APTOS 2019 database images plus augmented images, and the results are compared with popular image classification techniques, including SVM, KNN, DT, RF, LR, and NB. The flowchart of the proposed DL-based DR stages classification is shown in Fig. 3.

### DR severity stage classification using DCNN

The recognition of DR severity level can be achieved through CNN to extract features from a retinal color fundus image collected from the APTOS Kaggle database for identifying DR stages from class 0 to class 4. The deep

**Table 2** Summary of retinal images with a severity level of DR from the APTOS dataset

Severity level of DR	Number of image samples
Class 0 (normal)	1805
Class 1 (mild stage)	370
Class 2 (moderate stage)	999
Class 3 (severe stage)	193
Class 4 (proliferative stage)	295
Total	3662



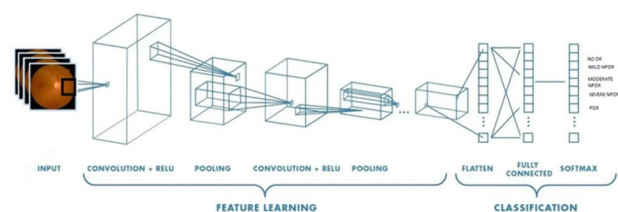
**Fig. 3** Flowchart for the proposed DR stage classification work

convolutional neural network consists of convolutional layers, a pooling layer, fully connected layers, and dense layers of output neurons with a softmax layer. This work applies the input RGB image to the first convolutional layer with a neuron size of  $224 \times 224 \times 3$ .

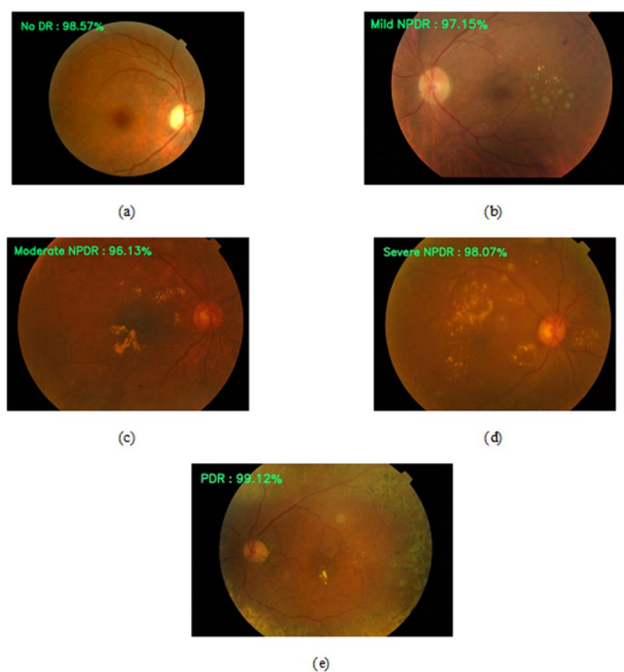
The proposed DCNN architecture is given in Fig. 4. Then, the input image is allowed through a stack of the convolutional layer with a “relu” activation function where the filter uses the small receptive field as  $3 \times 3$  with the convolution stride of one pixel. The padding is chosen to preserve the spatial resolution throughout the convolutional layers to extract 32 feature maps with the pool size of  $2 \times 2$ . “Relu” stands for rectified linear unit, which is used for learning non-linear real-world data. The second and third convolutional — relu with maximum pooling layers use the same  $3 \times 3$  kernel, resulting in 64 and 128 feature maps with the convolution stride of one pixel and the padding chosen as the same to produce the output of the same size as the input. The fully connected layer uses the SoftMax activation function for classifying the retinal images into five classes. In DCNN, the stochastic gradient descent (SGD) algorithm is used with a cross-entropy cost function for properly tuning hyper-parameters. The drop-out value is set as 0.5 in the convolutional layers and 0.25 in the fully connected layers to avoid overfitting issues during training.

### Implementation

In this work, the identification of severity stages has been completed through the training of the DCNN by assigning



**Fig. 4** The proposed DCNN architecture



**Fig. 5** Recognition results for severity stages of DR from retinal images: **a** no DR; **b** mild DR; **c** moderate DR; **d** severe DR; **e** PDR

a proper feature matrix and obtaining the target class based on the input class of the dataset. The images collected from the dataset with respective labels have been given as input to DCNN with a learning rate of 0.001 and trained for 50 epochs with a batch size of 32.

## Results and Discussion

In this section, the performance of the proposed method was evaluated and analyzed. However, the recognition rate in this proposed method is excellent, with the images used from the expanded image database for training and testing. The proposed algorithm classifies the severity stages of DR using the DCNN algorithm with an accuracy of 99%. Some of the images from the other datasets, such as DRIVE and images from the Google search engine, have been tested with the proposed model to evaluate the performance of our work. The classification of DR severity stages from class 0 to class 4 was successfully done in the proposed work. From these experimental results, it is visible that this algorithm can classify the severity stages of DR in real-time and is helpful for the early detection of DR with less cost and time.

**Table 3** Performance comparison of the various classifiers with different performance metrics

Name of the classifier	Classes	Precision	Recall	F1 score
DCNN	No DR	99	97	98
	Mild NPDR	100	99	99
	Moderate NPDR	98	100	98
	Severe NPDR	100	100	100
	PDR	99	99	99
SVM	No DR	90	93	91
	Mild NPDR	92	94	93
	Moderate NPDR	94	94	94
	Severe NPDR	100	97	99
	PDR	94	95	94
KNN	No DR	98	97	98
	Mild NPDR	100	100	100
	Moderate NPDR	100	94	97
	Severe NPDR	95	100	97
	PDR	96	98	97
DT	No DR	90	97	94
	Mild NPDR	90	100	95
	Moderate NPDR	86	91	88
	Severe NPDR	100	89	94
	PDR	81	70	75
NB	No DR	85	82	80
	Mild NPDR	71	86	77
	Moderate NPDR	84	68	75
	Severe NPDR	78	80	79
	PDR	88	75	81
RF	No DR	87	83	85
	Mild NPDR	76	80	78
	Moderate NPDR	78	71	74
	Severe NPDR	84	76	80
	PDR	80	74	77
LR	No DR	95	97	96
	Mild NPDR	96	94	95
	Moderate NPDR	95	96	95
	Severe NPDR	98	93	95
	PDR	97	96	96

## DR recognition and severity level prediction

The experimental part of this proposed work is divided into two main tasks. Task 1 involves the presence or absence of DR, which was identified. Another task is to predict the severity level of the DR stages using retinal images. For Task 1, the system must identify whether

**Table 4** Performance comparison of various classifiers concerning Accuracy

Model	Accuracy (%)
DCNN	98.5
K-NN	97.6
SVM	96.4
LR	95.5
DT	87.5
RF	84
NB	82.7

the patient has suffered with DR. In this work, the DR images from the dataset were labeled as the name of the disease's severity level. The remaining normal images are labeled as no DR and recognized successfully by our algorithm with a good detection score. Furthermore, identifying the severity level is most important to give better treatment. Hence, the classification of the severity level of DR for five different classes is carried out in the next task.

The severity level classification of DR was done on an expanded image dataset developed for this work with augmented images and the images from the APTOS Kaggle dataset with variations in illumination and brightness. The proposed work to detect the severity level of DR is performed on the laptop with the configurations of Intel Core i7-CPU@ 4.5 GHz and Google Colab. The programs for training and testing DR images were written in Python with OpenCV on the Microsoft Windows 10 operating system. The experimental tests were carried out on images listed under three categories in the Kaggle dataset such as test\_images, train\_images, and val\_images. Some of the results were also presented in Fig. 5 for the recognition of DR severity stages.

The severity stages of DR were tested with the proposed algorithm, and it was observed that the results were better when compared with other classification methods such as SVM, KNN, DT, LR, RF, and NB. Comparing the experimental results, the performance of NB was low among all the other methods. The proposed method can accurately detect the stages from class 0 to class 4. The NB method is unable to detect more abnormalities in the images. The experimental results indicated that the DCNN method is suitable for recognizing all DR stages with high detection accuracy.

The normal retinal images with no abnormalities were utilized during testing to meet the real-time challenges. To check the same, 20 images that did not contain DR were collected and tested to test the performance of the proposed algorithm from the Google search engine. The recognition method with DCNN found no abnormalities in these images, with a detection score of 98.57%.

From the experimental results, it is clear that our proposed DCNN model adequately detected the severity stages of DR with a detection score of 97.15% for mild NPDR, moderate NPDR at 96.13%, and severe NPDR at 98.07%, PDR at 99.12%. The results of this experiment further revealed that the DCNN model was the most suitable for detecting all stages of DR, mainly when there were minor abnormalities along with complex backgrounds in the retinal images. The RF method struggled to recognize the lesions well, but it performed better when compared with the NB method. These two methods needed help to recognize the severity stages positioned at various locations in the images and lesions that were grouped and close together with other objects.

Nevertheless, some lesions are misidentified in the retinal images for some severity levels, and the images vary in illumination in real-time. These issues will be resolved in the future by adding more retinal images in the database to improve learning with this vision-based technique, meet real-world challenges, and make this model more suitable for detecting more vision-threatening issues. The performance evaluation analyses were conducted, and the results were discussed in the forthcoming section.

## Performance evaluation analyses

### Classification of DR

The object recognition algorithm's performance depends on the classifier's performance. The analysis was evaluated by calculating the statistical indicators, including accuracy, precision, recall or sensitivity, and F1 score.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The performance comparison of the DCNN with other classifiers with performance metrics such as precision, recall, and F1 score measures is presented in Table 3. According to the classification metrics obtained from the validation step, the F1 score value of the DCNN for the classification of no DR, mild NPDR, moderate NPDR, severe NPDR, and PDR was 98, 99, 98, 100, and 99, and the overall F1 score achieved by the algorithm was 99% respectively. The results of the

proposed work showed that the DCNN classification algorithm is better than other classifiers such as SVM, KNN, DT, LR, RF, and NB for the classification of DR severity levels taken from the expanded retinal fundus image dataset.

Furthermore, the performance comparison related to accuracy is presented in Table 4. The experimental results witnessed that DCNN can recognize the major severity stages of DR. This underlying work acts as a base to detect the lesions present in the retinal fundus image in real time for the grading of DR in an automated manner.

The proposed DCNN algorithm performed well in the accuracy part, and it offers a good accuracy of 98.5% with no over-fitting issue by applying data augmentation and dropout techniques during training.

## Conclusion

This paper proposed a DR severity stage recognition and classification algorithm using the DL technique with the APTOS Kaggle dataset. In the proposed method, the severity stages were efficiently predicted from retinal fundus images. We evaluated the performance challenges by checking the algorithm with other dataset images, such as the DRIVE dataset and retinal fundus images from the Google search engine, and concluded that the proposed model can classify the severity level of DR successfully. Hence, the proposed method efficiently predicts DR with an accuracy of 98.5%. The detection of the lesions in the retinal fundus images using deep learning-based object detectors will be considered as future work to perform the grading of DR automatically without manual guidance.

**Acknowledgements** This research did not receive specific grants from public, commercial, or not-for-profit funding agencies. We are very thankful for Kaggle team for dataset availability for public use and we have used APTOS Kaggle dataset for our research work. We are very delightful for our management and principal for providing support to carry out this research work.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Ethical clearance** In this research work, we have utilized images from the Kaggle APTOS dataset and carried out the experimental work. And hence, the consent of patient, Ethical clearance are not required. In future, we will evaluate the performance of our work on real-time images.

## References

- Raman, RajivBhende, Pramod et al. Prevalence of diabetic retinopathy in India stratified by known and undiagnosed diabetes, urban–rural locations, and socioeconomic indices: results from SMART India population-based cross-sectional screening study. *Lancet Glob Health*. 2022;10(12): e1764 - 73.
- Bodapati JD, Veeranjanyulu N. Feature extraction and classification using deep convolutional neural networks. *J Cyber Secur Mobil*. 2019;8:261–76.
- Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Srivastava G. Deep neural networks to predict diabetic retinopathy. *J Ambient Intell Humaniz Comput*. 2020. <https://doi.org/10.1007/s12652-020-01963-7>.
- Das D, Biswas SK, Bandyopadhyay S. A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning. *Multimed Tools Appl*. 2022;81(18):25613–55. <https://doi.org/10.1007/s11042-022-12642-4>.
- Amin J, Sharif M, Yasmin M, Ali H, Fernandes SL. A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions. *J Comput Sci*. 2017;19:153–64. <https://doi.org/10.1016/j.jocs.2017.01.002>.
- Decenci ere E, Cazuguel G, Zhang X, Thibault G, Klein JC, Meyer F, Marcotegui B, Qu ellec G, Lamard M, Danno R, Elie D, Massin P, Viktor Z, Erginay A, Lay B, Chabouis A. TeleOphtha: machine learning and image processing methods for teleophthalmology. *IRBM*. 2013;34(2):196–203. <https://doi.org/10.1016/j.irbm.2013.01.010>.
- Nadeem MW, Goh HG, Hussain M, Liew SY, Andonovic I, Khan MA. Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions. *Sensors (Basel)*. 2022;22(18):6780. <https://doi.org/10.3390/s22186780>.
- Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AAS, Asari VK. A state-of-the-art survey on deep learning theory and architectures. *Electronics* 2019;8(3):1–66. <https://doi.org/10.3390/electronics8030292>.
- Kulkarni, Arun D. Deep convolution neural networks for image classification. (IJACSA) International J Adv Comput Sci Appl. 2022;13(6):2022. <http://hdl.handle.net/10950/4001>. Accessed 20 Mar 2024.
- Bhandari S, Pathak S, Amit S. A literature review of early—stage diabetic retinopathy detection using deep learning and evolutionary computing techniques. *Arch Comput Methods Eng*. 2022. <https://doi.org/10.1007/s11831-022-09816-6>.
- Qiao L, Zhu Y, Zhou H. Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. *IEEE Access*. 2020;8:104292–302.
- Saranya P, UmaMaheswari KM. Classification of different stages of diabetic retinopathy using convolutional neural networks. 2nd International Conference on Computation, Automation and Knowledge Management, (ICCA KM); 2021.
- Zhang W, et al. Knowledge-based systems automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowl Base Syst*. 2019;175:12–25. <https://doi.org/10.1016/j.knosys.2019.03.016>.
- Hemanth DJ, Deperlioglu O, Kose U. An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Comput Appl*. 2020;32(3):707–21. <https://doi.org/10.1007/s00521-018-03974-0>.
- APTOS 2019 Kaggle dataset, <https://www.kaggle.com/datasets/mariaherrero/aptos2019>. Accessed 20 Mar 2024.
- Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev*. 2023;56:12561–605. <https://doi.org/10.1007/s10462-023-10453-z>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.