International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Fine Tuned Pre-Trained Deep Neural Network for Automatic Detection of Diabetic Retinopathy Using Fundus Images

Chaitanya Singla[1], Ravneet Kaur*[2], Janpreet Singh[3], Nisha Nisha[4],Anshul Kumar Singh[5], Tejinder Singh[6]

Submitted: 27/04/2023 **Revised**: 25/06/2023 **Accepted**: 05/07/2023

Abstract: Diabetic Retinopathy (DR) is a visual condition that occurs because of chronic diabetes, causing damage to retinal blood vessels. It is considered a major cause of vision loss, affecting over 158 million people worldwide. Early detection and diagnosis can minimize the impact on vision and improve the quality of life for those with DR. Therefore, an automated DR detection method is essential. Deep learning models can automate classification and feature extraction, but developing such models from scratch requires a larger and well-annotated dataset. The main challenge in using deep learning for medical image analysis is the lack of annotated training datasets. To address this issue pre-trained deep learning networks are mainly used by the researchers. This study utilizes fine-tuned pre-trained deep learning techniques to derive attributes from fundus images, improving the accuracy of DR classification. To address data imbalances and insufficiencies, various techniques were used to enhance data for each DR class. The results of this study, evaluated using test data, show that our approach outperforms previous techniques in terms of accuracy.

Keywords: *Diabetic Retinopathy (DR), Deep Learning, Fine Tuned Deep Learning Network, Fundus Images, EyePACS*

1. Introduction

DR is an eye-related complication of diabetes that can cause blindness. It is a progressive disease that affects the retina, the light-sensitive tissue at the back of the eye and is considered the leading cause of blindness among working-age adults worldwide. The incidence of DR has increased by over a quarter in the last three years, with 40%-45% of diabetic patients at risk of developing DR. Diabetic Retinopathy (DR) can be broadly categorized into two main types: proliferative diabetic retinopathy (PDR) and non-proliferative diabetic retinopathy (NPDR). NPDR, in turn, is further divided into three types based on the severity of the disease. These include mild NPDR, moderate NPDR, and severe NPDR [1]. PDR occurs when abnormal blood vessels form on the surface of the retina, disrupting its normal blood supply. On the other hand, NPDR is characterized by weakened blood vessels without the formation of new ones.

Early indications of DR include microaneurysms (MA), hemorrhages (HA), and exudates (EX) [2]. In the initial phase of DR, the disease damages small blood vessels, leading to the formation of micro-aneurysms. MAs create small swellings around the edges of the tiny retinal vessels, and with the progress of the disease, hemorrhages can occur within the retina. HA can cause severe vision loss by damaging the retina.

As the disease progresses, the retinas weaken and begin to leak fluid, leading to the formation of lipid-based sediments called exudates. These exudates can be classified as either hard exudates (HEX) or soft exudates (SE), depending on their size and shape. The grading scheme of DR categories is described in Table 1.

Table 1. DR Categories

Grade	*Clinical Symptom*	*DR Stage*
0	No Symptoms	No DR
1	The presence of MA's that cause swelling on small retinal vessels presents in any quarter	Mild DR
2	HA, MA's	Moderate DR
3	As HA increases cause severe vision loss	Severe DR
4	HEX, SE, leaking fluid	Proliferate DR (PDR)

[1,2,5,6]*Chitkara University Institute of Engineering and Technology Chitkara University, Punjab, India*
[3]*School of Computer Science and Engineering Lovely Professional University, Punjab, India*
[4]*Deakin University*
** Corresponding Author Email: ravneetkaur.cse@chitkara.edu.in*

The manual screening process for DR is a tedious and time-consuming task and is often subject to the subjective interpretation of ophthalmologists [3]**.** Moreover, manual assessment of many DR patients fails to meet the screening demand, leading to inconsistent results among retina specialists [4]. Most DR patients are from low-income countries, where they face challenges in accessing retina specialists and other required facilities for DR diagnosis**.** Therefore, technology-based DR screening methods are necessary to overcome the challenges of manual diagnosis and assist ophthalmologists.

Fig. 1 shows the fundus images of each category of DR along with the image that does not have this disease.

Medical imaging employs image data to detect and identify abnormalities at an early stage. Computer-aided Diagnosis (CAD) devices enhance low-contrast images using various imaging techniques developed to identify prominent areas in the image.

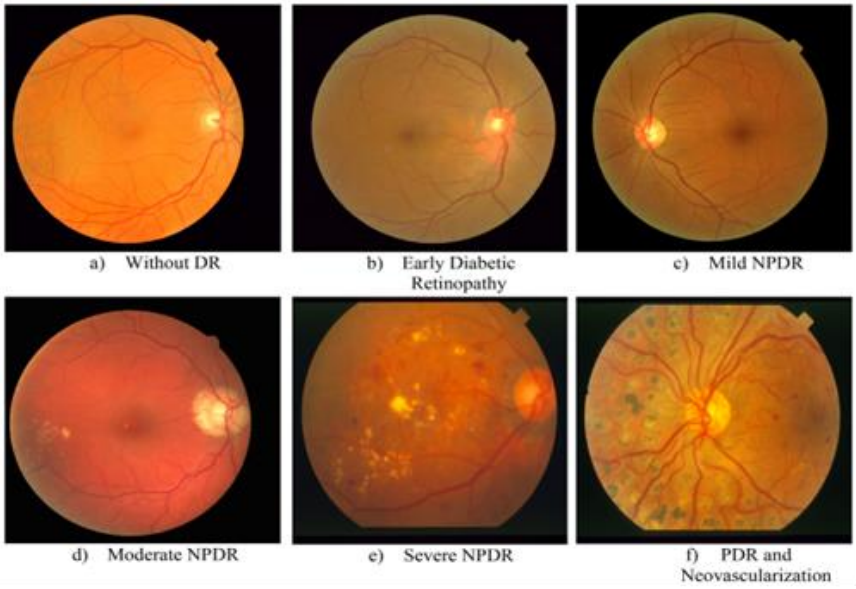


Fig. 1 DR Stages [5]

Such images provide valuable information for diagnosing diseases [6]. These systems are designed to provide experts in the field with a second opinion, thus improving overall diagnostic performance and assisting doctors in detecting health issues early in the disease course. These systems are designed to improve the accuracy of disease diagnosis and simplify the processing of image information. Clinics also play a role in this area, and computer-aided algorithms are expected to improve diagnostic accuracy for better performance. The CAD system is also useful for fundus analysis, and with frequent screening and timely intervention, the risk of developing serious eye problems can be reduced. This research aims to provide a CAD algorithm that will assist ophthalmologists in distinguishing different types of HR and DR with an additional judgement [7]. A pre-trained network consists of multiple filters, with each layer having a variety of characteristics. Building an autonomous classifier with a large number of neurons and weights is a difficult task. Therefore, a system that can extract the most features and has a flexible classification structure is essential.

The main aim of this research is to propose an effective method for optometrists to provide a second opinion in categorizing early-stage DR. The research aims are as follows:

1. Applying Fine-Tuned Deep Neural Networks on datasets to classify DR.
2. Implementing different pre-processing techniques, including extracting green-channel images using RGB Image, CLAHE for visibility enhancement, and Morphological operations to determine DR position.
3. Performing Data Augmentation for each DR grade to address the issue of insufficient data and ensure balanced data.
4. Developing an efficient system capable of classifying each class. The system is evaluated using EYEPACS data based on various performance indicators, including accuracy, recall, precision, F1-score, and accuracy.

The research conducted in this paper is structured as follows: Section 2 presents a literature review. Section 3 outlines the research proposal. Section 4 covers Evaluation Metrics, while Section 5 includes the conclusion and future research.

2. Literature Review

Ophthalmologists may take more time while doing manual diagnosis of the patients who may be infected with the DR disease. While diagnosis it might prone to misdiagnosis. The use of computer-aided diagnosis (CAD) can help to prevent misdiagnosis and reduce the time, cost, and energy involved. Deep learning (DL) has gained popularity in various industries over the last decade, including medical imaging analysis. DL outperforms other algorithms for image analysis by accurately identifying features of input data to aid in classification or segmentation. Unlike other methods, DL does not require the creation of custom-designed features, but it requires a huge labelled data to train.

Convolutional neural network (CNN) is an efficient and effective method of DL for image processing, often used [8], [9].

In [10] authors applied CNN model on the two fundus images datasets and achieved 96% in one and 75.09% in another dataset. A deep CNN was used to identify and classify DR in five different stages, achieving 94.44% accuracy when applied to the Kaggle database [11].

A multi-tasking deep neural model was developed to determine the severity of DR, composed of a classification model and a regression model, each having its loss function. The regression and classification models were used together to explain the dependence between the severity stages. The Dense Net modification was employed to carry out this strategy, achieving 85% and 82% accuracy when tested using APTOS and EYEPACS datasets, respectively. To evaluate the research, the Xception network using transfer learning was also designed [12].

3. Proposed Methodology

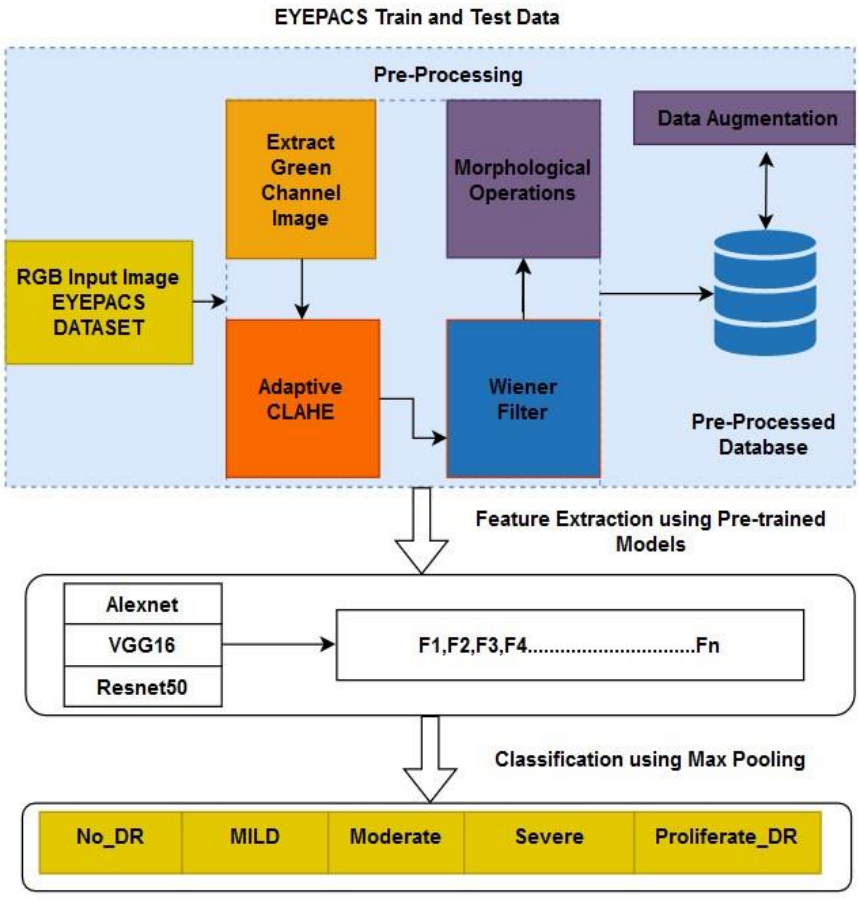


Fig. 2. Proposed Methodology

Methodology is the process that helps to analyze the data, and classify, or categorize the data into different categories. The proposed methodology consists of three processes that include pre-processing of the image for image enhancement and data augmentation to remove the data deficiency problem. Taking the balanced dataset as the input extracting the features and classifying them into five DR categories is the last step. The extraction of features and classification part is done using transfer learning. Fig. 2. shows the proposed methodology.

3.1 Image Pre-Processing

Retina fundus images were obtained using the publicly available dataset EyePACS, which was accessed through

Kaggle.com on May 8, 2023. The EyePACS dataset is diverse and includes images of different sizes and illumination levels. To ensure consistency, all images were resized to a common size, and the green channel was extracted from the RGB images. Fig. 3 illustrates that the green channel contains more retinal information compared to the blue or red channels.

The contrast-limited adaptive histogram equalization (CLAHE) was introduced to enhance the quality of the retinal image, including the enhancement of tiny regions. Subsequently, the Wiener filter and morphological operation were applied to the green channel image to further improve its quality.

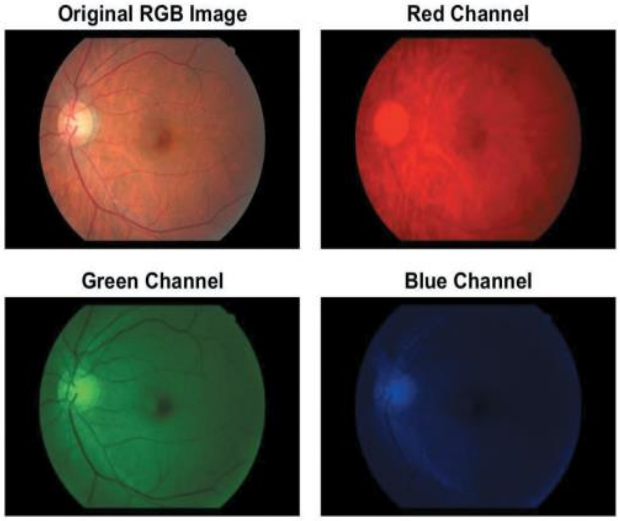
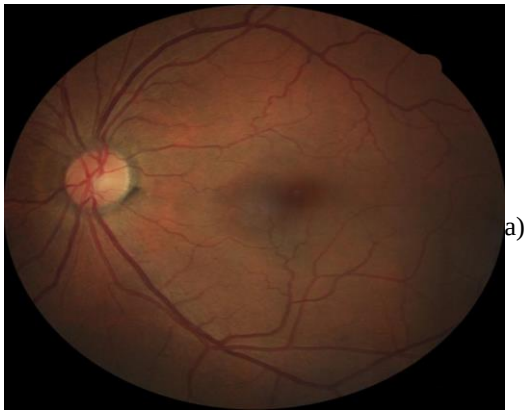
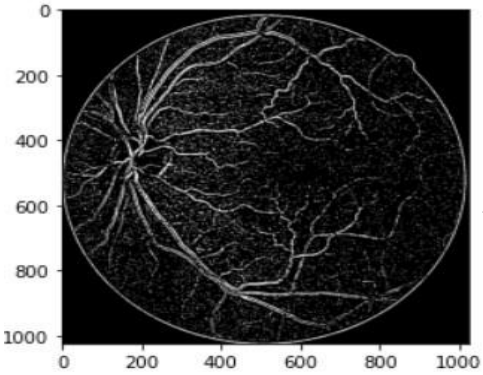


Fig. 3. RGB Channels of Original Fundus Image

Fig. 4a shows the original image, while Fig. 4b displays the processed image.


a)



b)

Fig. 4. a) Original Image b) Preprocessed Image

3.2 Data Augmentation

The size of the training dataset is one of the critical factors that affect the effectiveness of DL models. Therefore, it's crucial to use a larger dataset for training deep-learning architectures to avoid issues with overfitting and generalization. The dataset was highly imbalanced, as shown in Fig. 5, with most of the images in grade 0. The imbalanced dataset ratio was (36:3:7: 1:1), which could lead to inaccurate classification. To enhance the retinal database across different scales, data augmentation was performed. Since grade 0 data was available in large quantities, no augmentation was applied. Various augmentation methods were used to balance the data across all four grades. Initially, the training dataset size was 35,126 photos, including all grades, without augmentation. After augmentation, the total dataset size was increased to 127,505, including all grades.

3.3 Fine-Tuned Pre-Trained Models

While deep learning algorithms can be effective in solving various classification problems, the primary challenge in medical image classification is the lack of labeled data. Transfer Learning is a widely used technique to address the shortage of annotated data by leveraging pre-trained deep convolutional neural networks for a similar task. It can reduce learning overhead and allow training on a smaller dataset.

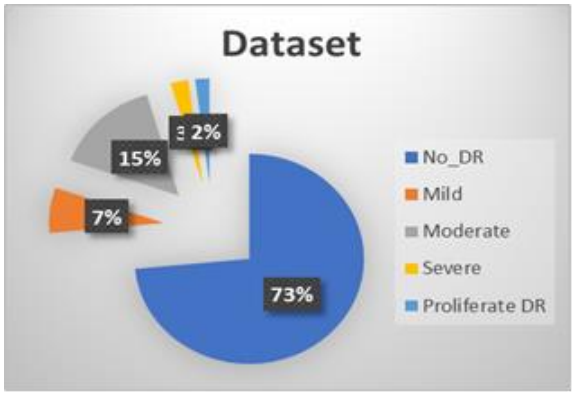


Fig. 5. Dataset (Category Percentage without

augmentation)

The use of transfer learning to improve the accuracy of annotated data for medical image classification using deep learning is a highly debated topic. In this study, we utilized pre-trained models that are AlexNet, VGG16, and ResNet50.

The initial weights of the model were frozen before fine-tuning, during which the last three layers (which were fully connected) were added, including the final output layer. The model layers were fine-tuned using the augmented dataset. The final output layer had a numeral of neurons equal to the total number of classes. Once the fine-tuning of the model layers was complete, the models were trained.

4. Performance Matrices and Result Analysis

The evaluation of all the three trained models is done based on the performance matrices such as accuracy, precision, Recall and F1-score. For all the three models the confusion Metrix is also created to show the class wise performance [13].

Precision: The proportion of numeral of true positive (TP) predictions and the summation of TP and false positives (FP) predictions. A high precision score indicates that the model has a good ability to accurately identify positive instances and avoid falsely labelling negative instances as positive. The precision calculated class wise using the above methodology is shown in the Fig. 6.

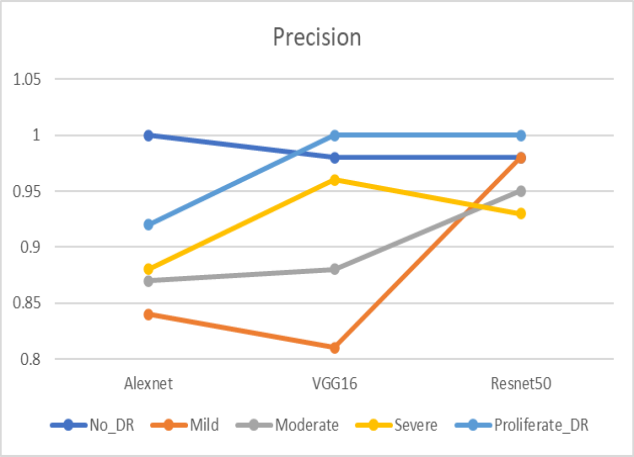


Fig. 6. Precision

Recall: The proportion of numeral of true positive (TP) predictions and the summation of TP and false negative (FN) predictions. A high recall score indicates that the model is skilled at identifying positive instances, but may also incorrectly classify some negative instances as positive. The precision calculated class wise using the above methodology is shown in the Fig. 7.

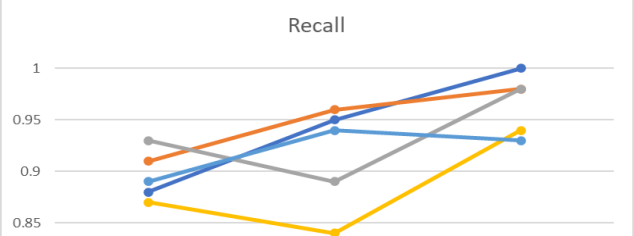


Fig. 7. Recall

F1-score: The proportion of twice of multiplication of precision and recall by the summation of precision and recall is called F1-score. A perfect F1 score is 1.0 and a lower score indicates poorer performance. A high F1-score indicates that the model has both good precision and recall, meaning it can accurately identify positive instances while avoiding false positives and negatives. The F1-score calculated class wise using the above methodology is shown in the Fig. 8.

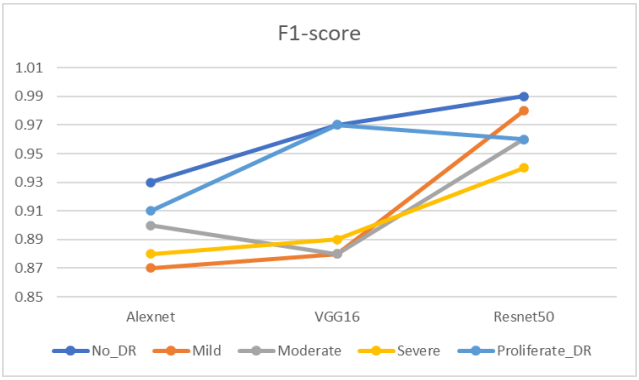
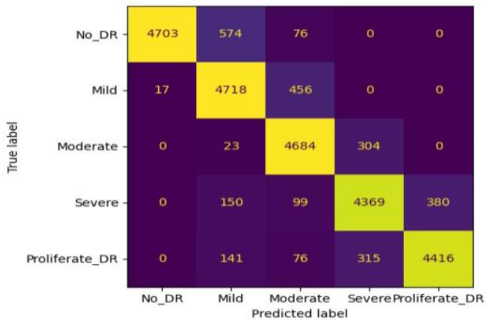


Fig.8. F1-score

The confusion matrix is also created using the three mentioned fine-tuned algorithms applied on the dataset. The confusion matrix and the results of the algorithms is shown in below figures 9,10,11. Figure 9 shows the results and the confusion matrix for the AlexNet model. Similarly figure 10 and figure 11 show the results and confusion matrix for VGG16 and ResNet50 model.



(a)

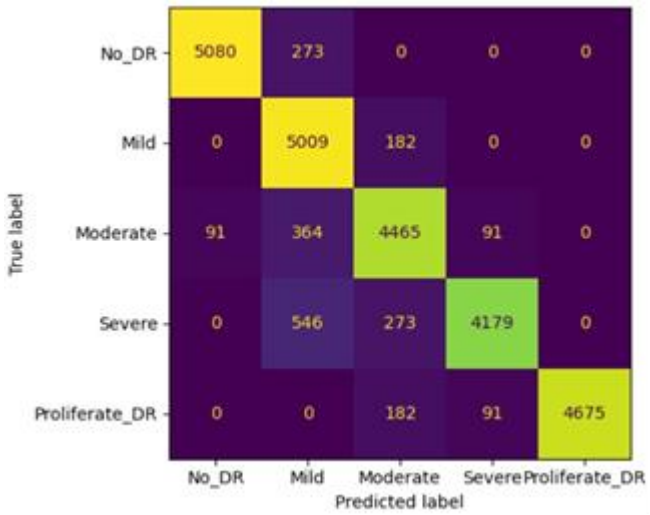
eering

	precision	recall	f1-score	support
No_DR	1.00	0.88	0.93	5353
Mild	0.84	0.91	0.87	5191
Moderate	0.87	0.93	0.90	5011
Severe	0.88	0.87	0.88	4998
Proliferate_DR	0.92	0.89	0.91	4948

b)

Fig. 9. AlexNet Confusion Matrix and Results

Fig. 9 shows that the accuracy of the AlexNet model after applied on the dataset comes out to be 90%. The value of the precision is 1 in case of No_DR stage.



a)

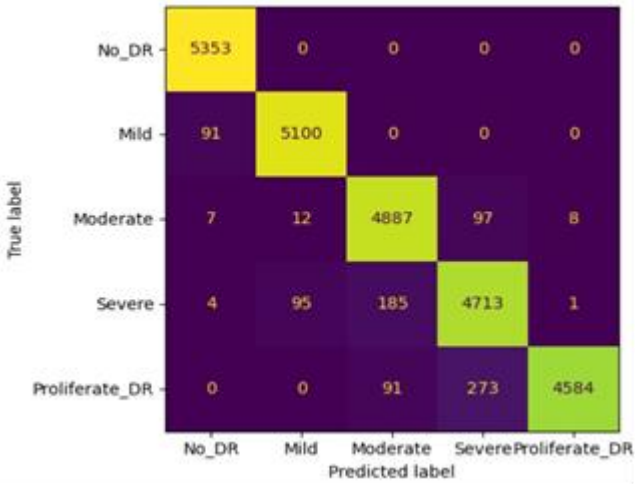
	precision	recall	f1-score	support
No_DR	0.98	0.95	0.97	5353
Mild	0.81	0.96	0.88	5191
Moderate	0.88	0.89	0.88	5011
Severe	0.96	0.84	0.89	4998
Proliferate_DR	1.00	0.94	0.97	4948
accuracy			0.92	25501
macro avg	0.92	0.92	0.92	25501
weighted avg	0.92	0.92	0.92	25501

b)

Fig. 10. VGG16 Confusion Matrix and Results

Fig. 10 shows that the accuracy of the VGG16 model after applied on the dataset comes out to be 92%. The value of the precision is 1 in the case of Proliferate_DR stage.

Fig. 11 shows that the accuracy of the ResNet50 model after applied on the dataset comes out to be 97%. The value of the precision is 1 in the case of Proliferate_DR stage. The fine-tuned pre-trained models demonstrated excellent performance. The training and validation accuracy rates continuously improved with each epoch.



a)

	precision	recall	f1-score	support
No_DR	0.98	1.00	0.99	5353
Mild	0.98	0.98	0.98	5191
Moderate	0.95	0.98	0.96	5011
Severe	0.93	0.94	0.94	4998
Proliferate_DR	1.00	0.93	0.96	4948
accuracy			0.97	25501
macro avg	0.97	0.97	0.97	25501
weighted avg	0.97	0.97	0.97	25501

b)

Fig. 11. ResNet50 Confusion Matrix and Results

The loss curve also shows a decrease in both training and validation loss with each epoch. A comparative analysis was conducted with state-of-the-art algorithms, and the results are presented in Table 2. The ResNet50 algorithm outperformed other algorithms in the automatic detection of DR stages.

Table 2. Result Comparison with other state of art-techniques

Author/ Year	Model	Accuracy
[14], 2018	CNN	85.30 %
[15], 2018	CNN	78.70 %
[16],2018	CNN	86.10 %
[17], 2018	DCNN	86.04 %
[18], 2018	CNN	84.36 %
[19], 2022	CNN	90.40 %
[20], 2020	CNN	82.20 %
[21], 2021	CNN	87.06 %
[22],2022	Transfer Learning	96.60 %
Proposed	Fine Tuned Transfer Learning	97.0

5. Conclusion and Future Work

We have proposed a method to automate the recognition and diagnosis of diabetic retinopathy using the concept of transfer learning in this article. Various pre-processing techniques are employed to enhance the quality of images. To overcome the issue of insufficient data, data augmentation techniques are utilized. Following pre-processing and data augmentation, refined pre-trained models are used to classify DR levels. The features that appear in fundus images are extracted using three different architectures and then combined using transfer learning to improve the classification performance. Additionally, we applied different data augmentation techniques to each level of DR separately to produce a balanced dataset and enhance the effectiveness of the design. Finally, the study results were evaluated against various deep learning algorithms, and the findings were compared to conventional techniques. The technique showed excellent performance in various statistical accuracy measures. We plan to combine hand-engineered features with CNN to further improve the classification accuracy.

References

[1] Mahendran G, Dhanasekaran R, and Narmadha Devi K.N, "Identification of exudates for Diabetic Retinopathy based on morphological process and PNN classifier," *2014 International Conference on Communication and Signal Processing*, Apr. 2014, doi: https://doi.org/10.1109/iccsp.2014.6950022.

[2] C. Agurto *et al.*, "Multiscale AM-FM Methods for Diabetic Retinopathy Lesion Detection," *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 502–512, Feb. 2010, doi: https://doi.org/10.1109/tmi.2009.2037146.

[3] K. Jin *et al.*, "Automatic detection of non-perfusion areas in diabetic macular edema from fundus fluorescein angiography for decision making using deep learning," *Scientific Reports*, vol. 10, no. 1, p. 15138, Sep. 2020, doi: https://doi.org/10.1038/s41598-020-71622-6.

[4] Idris, L. Sellahewa, C. Simpson, P. Maharajan, and J. Duffy, "Grader agreement, and sensitivity and specificity of digital photography in a community optometry-based diabetic eye screening program," *Clinical Ophthalmology*, p. 1345, Jul. 2014, doi: https://doi.org/10.2147/opth.s61483.

[5] D. Nagpal, S. N. Panda, M. Malarvel, P. A. Pattanaik, and M. Zubair Khan, "A review of diabetic retinopathy: Datasets, approaches, evaluation metrics and future trends," *Journal of King Saud University - Computer and Information Sciences*, Jun. 2021, doi: https://doi.org/10.1016/j.jksuci.2021.06.006.

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

[7] L. Deng, "Deep Learning: Methods and Applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 3–4, pp. 197–387, 2014, doi: https://doi.org/10.1561/2000000039.

[8] M. Bakator and D. Radosav, "Deep Learning and Medical Diagnosis: A Review of Literature," *Multimodal Technologies and Interaction*, vol. 2, no. 3, p. 47, Aug. 2018, doi: https://doi.org/10.3390/mti2030047.

[9] G. Litjens *et al.*, "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: https://doi.org/10.1016/j.media.2017.07.005.

[10] M. K. Yaqoob, S. F. Ali, M. Bilal, M. S. Hanif, and U. M. Al-Saggaf, "ResNet Based Deep Features and Random Forest Classifier for Diabetic Retinopathy Detection," *Sensors*, vol. 21, no. 11, p. 3883, Jun. 2021, doi: https://doi.org/10.3390/s21113883.

[11] D. Nagpal, N. Alsubaie, B. O. Soufiene, M. S. Alqahtani, M. Abbas, and H. M. Almohiy, "Automatic Detection of Diabetic Hypertensive Retinopathy in Fundus Images Using Transfer Learning," *Applied Sciences*, vol. 13, no. 8, p. 4695, Apr. 2023, doi: https://doi.org/10.3390/app13084695.

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

[13] Moudgil, S. Singh, V. Gautam, S. Rani, and S. H. Shah, "Handwritten devanagari manuscript characters recognition using capsnet," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 47–54, Jun. 2023, doi: https://doi.org/10.1016/j.ijcce.2023.02.001.

[14] Rakhlin, "Diabetic Retinopathy detection through integration of Deep Learning classification framework," BioRxiv, Nov. 2017, doi: https://doi.org/10.1101/225508.

[15] K. Chang *et al.*, "Distributed deep learning networks among institutions for medical imaging," *Journal of the American Medical Informatics Association*, vol. 25, no. 8, pp. 945–954, Mar. 2018, doi: https://doi.org/10.1093/jamia/ocy017.

[16] G.-M. Lin *et al.*, "Transforming Retinal Photographs to Entropy Images in Deep Learning to Improve Automated Detection for Diabetic Retinopathy," *Journal of Ophthalmology*, vol. 2018, p. e2159702, Sep. 2018, doi: https://doi.org/10.1155/2018/2159702.

[17] Z. Li *et al.*, "An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs," *Diabetes Care*, vol. 41, no. 12, pp. 2509–2516, Oct. 2018, doi: https://doi.org/10.2337/dc18-0147.

[18] S. Seth and B. Agarwal, "A hybrid deep learning model for detecting diabetic retinopathy," *Journal of Statistics and Management Systems*, vol. 21, no. 4, pp. 569–574, Jun. 2018, doi: https://doi.org/10.1080/09720510.2018.1466965.

[19] M. N. Pal and M. Banerjee, "Retinal vessel segmentation using a strip wise classification approach with grid search-based parameter selection," *International Journal of Computational Vision and Robotics*, vol. 12, no. 2, p. 194, 2022, doi: https://doi.org/10.1504/ijcvr.2022.121187.

[20] N. Tamim, M. Elshrkawey, G. Abdel Azim, and H. Nassar, "Retinal Blood Vessel Segmentation Using Hybrid Features and Multi-Layer Perceptron Neural Networks," *Symmetry*, vol. 12, no. 6, p. 894, Jun. 2020, doi: https://doi.org/10.3390/sym12060894.

[21] E. Uysal and G. E. Güraksin, "Computer-aided retinal vessel segmentation in retinal images: convolutional neural networks," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3505–3528, Sep. 2020, doi: https://doi.org/10.1007/s11042-020-09372-w.

[22] M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, "Transfer Learning-Based Model for Diabetic Retinopathy Diagnosis Using Retinal Images," *Brain Sciences*, vol. 12, no. 5, p. 535, Apr. 2022, doi: https://doi.org/10.3390/brainsci12050535.

[23] Dinakar, J. R. ., & S., V. . (2023). Real-Time Streaming Analytics using Big Data Paradigm and Predictive Modelling based on Deep Learning . International Journal on Recent and Innovation Trends in Computing and Communication, 11(4s), 161–165. https://doi.org/10.17762/ijritcc.v11i4s.6323

[24] Prof. Amruta Bijwar. (2016). Design and Analysis of High Speed Low Power Hybrid Adder Using Transmission Gates. International Journal of New Practices in Management and Engineering, 5(03), 07 - 12. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/46