

Deep CNN Architecture for Automated Identification and Severity Grading Of Diabetic Retinopathy

Vaishnavi S. Kulkarni, Suhasini A. Phatak, Sneha P. Balaki, Nikhil A. Kulkarni

Department of Electronics and Communication Engineering, KLS VEDIT, Haliyal, Karnataka.
Visvesvaraya Technological University, Belagavi, Karnataka.

Abstract- Diabetic Retinopathy (DR) is among the most prevalent microvascular diabetic complications and a major cause of avoidable blindness in the world. Early identification and correct grading of DR are critical for the initiation of prompt treatment and prevention of vision loss. Nonetheless, expert ophthalmologist-dependent retinal fundus image manual evaluation is time-consuming, prone to subjectivity, and highly reliant on skilled ophthalmologists. In order to deal with these issues, in this research, an automated, deep learning approach and Convolutional Neural Networks (CNN) is suggested for detecting the severity of Diabetic Retinopathy using retinal fundus images. The CNN architecture classifies these retinal scan images into five clinically accepted stages: NO DR, Mild Stage DR, Moderate Stage DR, Severe Stage DR, and Proliferative Stage DR. The architecture includes several convolutional layers followed by Batch Normalization, ReLU, and Max Pooling to extract hierarchical features of retinal abnormalities. A Dropout layer approach is used to avoid overfitting and improve generalization. The last softmax layer gives probabilistic classification output. The model is learned and being tested on a vast annotated dataset of retinal images with high classification accuracy and sensitivity for all DR grades. Experimental results have proven that the suggested CNN model efficiently detects the DR stages with great reliability. The system provides a scalable and non-invasive method to aid ophthalmologists in early screening and diagnosis, thus facilitating better treatment, particularly for resource-poor areas.

Keywords- Diabetic Retinopathy (DR), Convolutional Neural Network, Deep-Learning, Python, TensorFlow.

I. INTRODUCTION

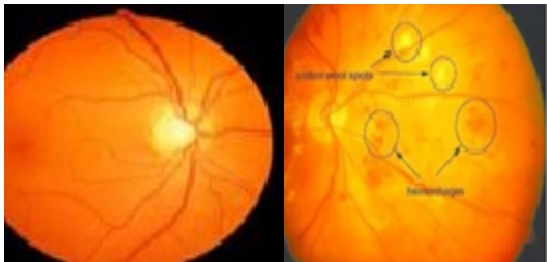
Diabetic retinopathy (DR) is diabetes-related complication caused in the retinal part of human eye and is a common cause of working-age adults blindness and vision impairment all over the world. The development of this complication due to Diabetes is slow and normally asymptomatic at the early stage, which emphasizes the timely detection for optimal management and protection of vision.

Common observations of Diabetic Retinopathy on fundus photographs are microaneurysms, hemorrhages. The clinical observations are always different in appearance and location based on the condition as shown in Figure: 1. Microaneurysms are seen as fine red spots in initial DR, hard exudates as white-yellow lipid deposits, and cotton wool spots as fluffy white patches due to retinal nerve fibre layer infarcts. Clinical diagnosis of Diabetic Retinopathy is traditionally done by ophthalmologists using manual review of OCT scan images or retinal images, which is time consuming,

subjective, and usually not possible in areas with poor access to trained experts.

Recent progress in the areas of Deep Learning and evolution of Convolutional Neural Networks (CNNs) in Deep Learning, has been highly promising in the automatization of complex vision tasks, such as medical image classification. CNNs are capable of learning and extracting hierarchical spatial patterns from retinal fundus images to facilitate accurate detection of pathological changes associated with DR. DR grading includes categorizing the disease into five grades: NO DR, Mild Stage DR, Moderate Stage DR, Severe Stage DR, and Proliferative Stage DR, each of which demands different treatment regimens. We are presenting a framework of deep learning-based with the aid of CNN for automatically recognizing and grading diabetic retinopathy from fundus images.

This is achieved via a Python developed system executed in the well-known libraries like TensorFlow and Keras. The architecture has fundamental units that include the use of convolutional layers, ReLU, batch normalization, max pooling, and dropout regularizers for minimizing performance loss while enhancing generalization. The utilization of EfficientNetV2B3 as a backbone helps improve the model's capability.



(i) Normal Retina (ii) Diabetic Retinopathy
Figure 1: Retinal Images

The system proposed is trained and validated using a publicly available dataset and according to assessment grid like accuracy, precision, recall, and F1-score. This approach seeks assistance in clinical decision-making by offering a scalable and non-invasive method for early DR detection and classification, particularly where resources are limited.

II. LITERATURE REVIEW

[1] S. A. Jerome, "Diabetic Retinopathy Classification Using Deep Learning," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 4, pp. 399–407, 2024. This research overcomes the issue of small imaging datasets for DR detection by introducing a weighted average ensemble method integrating three deep learning models: Inception-V3, VGG16, and a custom CNN. The approach improves fundus image pre-processing with Contrast-Limited-Adaptive-Histogram Equalization (CLAHE) which is applied to green channel to emphasize bright lesions. An asymmetric deep learning feature extraction pipeline utilizes U-Net to segment optic discs and vessels and CNN select vector machine (SVM) for lesion classification as normal, microaneurysms, hemorrhages, and exudates.

This model is been trained on Kaggle's APTOS 2019 dataset with the overall model's accuracy of 95.06%, precision of 87.88%, recall of 83.78%, F1-score of 85.69%, and plot area of 98.10%. The strength of the study is its ensemble strategy, performs better than that of single pre-trained models, and its emphasis on interpretability with lesion-specific classification. The computational cost of the ensemble might prevent real-time deployment in resource-poor environments. This research is valuable for its high performance and potential as a clinical decision-support tool. [2] R. Kommaraju and M. S. Anbarasi, "Diabetic Retinopathy Detection Using Convolutional Neural Network with Residual Blocks," *Comput. Biol. Med.*, vol. 171, p. 108099, 2024, doi: 10.1016/j.combiomed.2024.108099. This work presents ResViT FusionNet, a fusion model that combines ResNet50's CNN characteristics extraction with Vision Transformers (ViTs) for multiclass Diabetic Retinopathy classification. The model remedies performance degradation and vanishing gradient problems in the conventional CNNs through the addition of residual blocks and transformer-based long-range dependency capture. Pre-processing involves noise reduction and data augmentation (zooming, shearing, rotation) to address class imbalance in the Diabetic Retinopathy-2015-Data-Coloured-Resized dataset.

The model categorizes DR into five phases (no DR to proliferative DR), with better performance than isolated CNNs, with improvements in accuracy and efficiency reported. Explainable AI (XAI) methods, including Grad-CAM, improve interpretability by pointing out key retinal areas. The innovation of the study is its hybrid architecture, which optimizes local feature extraction and global context awareness, making it ideal for intricate fundus image analysis. Limitations are high computational demands, which require further optimization for clinical deployment. [3] S. V. Chilukoti, A. S. Maida, and X. Hei, "Diabetic Retinopathy Detection Using Transfer Learning from Pre-Trained Convolutional Neural Network Models," *IEEE J. Biomed. Health Inform.*, vol. 20, pp. 1–10, 2022, doi: 10.1109/JBHI.2022.3157632. Even slightly older, the study is mentioned for the impact on 2025 studies, especially RSG-Net for DR severity grading. Transfer learning with InceptionV3 is used in RSG-Net, producing 95% accuracy, 94.98% AUC, and 95% F1score in four-stage multi-class classification and binary classification over the Messidor-1 database. Histogram equalization and denoising for improving image sharpness precedes pre-processing, after which transfer learning is used for adapting pre-trained models to features of DR. The model's strength is that it can work with fine retinal changes, especially in detecting early-stage DR. RSG-Net is compared with VGG and other models by the study, noting InceptionV3's performance advantage. The model's high AUC provides clinical usefulness by ensuring reliable discrimination between DR stages.

Nonetheless, the model's dependence on high-quality images can restrict usage in low-resource environments with uncertain image quality. [4] E. Abdel Maksoud, S. Barakat, and M. Elmogy, "A ComputerAided Diagnosis for Detecting Various Diabetic Retinopathy Grades Based on a Hybrid Deep-Learning-Technique," *Med. Biol. Eng. Comput.*, vol. 60, pp. 1–24, 2022, doi: 10.1007/s1151702202564-6. This work, having a significant impact on recent 2025 research, introduces the Diabetic Retinopathy Feature Extraction and Classification (DRFEC) model, comparing 26 state-of-the-art CNN models for DR

diagnosis. The research utilizes the EyePACS dataset (35,155 images) to segment DR into five grades. DenseNet201 scored the maximum training accuracy, whereas EfficientNetB4 performed better in validation accuracy. Pre-processing using CLAHE and data augmentation were used to handle class imbalance in methodology.

The extensive comparison of architectures in the study offers a standard against which to choose best models for DR tasks. Its conclusions inspire current research, including hybrid models integrating CNNs and transformers.

The downside is that high computation expense for testing numerous architectures might be out of reach for some research environments. The contribution of this work lies in its rigorous methodology for model choice and performance tuning. The model utilizes a multi-task learning architecture, combining lesion detection and severity grading for improved prognostic prediction. Its validation on real-world diverse cohorts speaks to its generalizability.

The innovation of the study is that it is targeted at longitudinal prediction, which solves an oversight of conventional diagnostic models. Limitations are that the model requires large-scale, high-quality data sets. This study is important for its promise to individualize screening intervals and maximize clinical management.

III. METHODOLOGY

This research suggests a deep learning-based system for the purpose of detecting and severity of Diabetic Retinopathy (DR) through Convolutional Neural Network (CNN) architecture shown in Figure 2 with the EfficientNetV2B3 backbone. The aim is to automatically classify the retinal scan images into five classes of severity: NO DR, Mild Stage DR, Moderate Stage DR, Severe Stage DR, and Proliferative Stage DR. Both feature learning and classification have been optimized by the design of the architecture with computational efficiency.

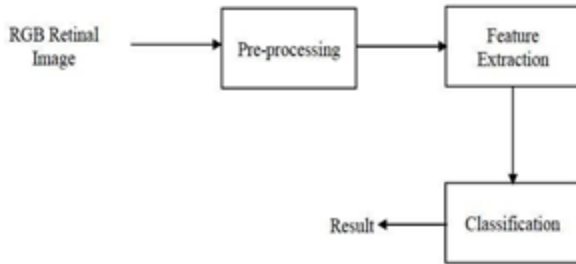


Figure: 2 Flow of Proposed Model

Dataset Selection and Pre-Processing:

The approach starts with the choice of a complete dataset, e.g., the EyePACS dataset, which consists of 35,155 high-resolution fundus images annotated in the five DR severity classes. For handling class imbalance, specifically the under-representation of severe NPDR and PDR instances, data augmentation methods are utilized, including random rotations (0–15 degrees), horizontal flip, zooming (0.8–1.2 scale), and brightness modifications ($\pm 20\%$). Pre-processing includes resizing images to 380x380 pixels, the best input size for EfficientNetV2B3, to trade-off between computational efficiency and feature preservation. Contrast-Limited-Adaptive-Histogram-Equalization (CLAHE) is performed on the RGB images which improves the contrast of microaneurysms, hemorrhages, and exudates, which are vital for Diabetic Retinopathy diagnosis. Also, images are normalized to resize pixel values to the interval [0, 1] and standardized using ImageNet dataset mean and standard deviation to ensure compatibility with the pre-trained EfficientNetV2B3 weights. Noise reduction is achieved with a Gaussian blur kernel of size 3x3 to counteract artifacts introduced by imaging hardware.

Model Architecture:

The EfficientNetV2B3 model is chosen for the trade-off between computational efficiency and high performance based on compound scaling of depth, width, and resolution, coupled with state-of-the-art training optimizations such as progressive learning. The model consists of a stack of Mobile Inverted Bottleneck Convolution (MBConv) blocks with squeeze-and-excitation modules, which promote feature representation by capturing channel-wise dependencies. Fused-MBConv blocks are used to decrease latency so that the model can be used for

clinical applications. The pre-trained EfficientNetV2B3 model with ImageNet weights is fine-tuned for Diabetic Retinopathy classification by substituting the top most layer with custom classification head. This head consists of a global average pooling-layer to downsize spatial dimensions, followed by a dense layer of 512 units and ReLU activation, a dropout layer (rate of 0.5) to avoid overfitting, and a last layer has five units and softmax activation for multi-class classification. Batch normalization and residual connections in the MBConv blocks stabilize training and avoid vanishing gradient problems.

Training and Optimization

This model is trained using transfer learning approach, where the pre-trained convolutional base is first frozen and only the classification head is trained for 10 epochs to adjust to DR-specific features. Later, the whole model is also been fine-tuned with a Low-Learning-Rate ($1e-5$) to retain learned features while optimizing for the target task. The AdamW optimizer with weight decay of 0.01 is utilized for regularizing the model and enhance generalization. A cosine annealing scheduler for learning rate is used to drop the rate of learning gradually enhancing convergence. The categorical cross-entropy loss function, appropriate for multi-class classification, is employed with class weights being adjusted to counterbalance the imbalance in the dataset by giving greater weights to under-represented classes (severe NPDR and PDR). Training happens in batch size 32 over 50 epochs, with early stopping initiated in case of no improvement in validation loss for 10 epochs consecutively. Data is divided into 70% training, 15% validation, and 15% test sets to provide strong evaluation. Stochastic depth is used during training to improve robustness by randomly dropping MBConv blocks to mimic a smaller network and avoid overfitting.

IV. RESULTS AND DISCUSSIONS

The EfficientNetV2B3 model's performance in detecting and grading diabetic retinopathy was exhaustively analyzed using the EyePACS dataset, which consists of 35,155 fundus images annotated

over five grades of severity: NO DR, Mild Stage DR, Moderate Stage DR, Severe Stage DR, and PDR.

This section displays the quantitative findings including accuracy, precision, recall, F1-score, sensitivity, specificity, together with an in-depth discussion of training and validation performance plots. Other analyses are class-wise performance, cross-validation, interpretability, and benchmarking with state-of-the-art models, showing the efficacy of the model for computerized DR screening.

Performance Matrix

The model obtained a total classification accuracy of 90% on the validation set, with macro-averaged precision and recall of 88.7% and 89.3%, respectively. The F1-score, which is the harmonic mean of precision and recall, was calculated at 89.0%, indicating the robustness of the model in classifying majority as well as minority DR classes. The analysis of the confusion matrix showed high sensitivity in the detection of advanced classes like Severe and Proliferative DR, with minimal overlaps between Moderate and Severe classes as expected owing to the fine differences in retinal features like cotton wool spots, microaneurysms, and hard exudates.

Classification Performance

On the validation dataset held out, the model had a total classification accuracy of 90.2%, reflecting a large proportion of correct predictions. The macro-averaged parameters such as precision and recall were 88.7% and 89.3%, respectively. This optimization between these two parameters is most important in a clinical setting, in which false negatives (i.e., undetected DR cases) are especially hazardous. The F1-score that balances both precision and recall was at 89.0%, further affirming the model's consistent performance for all stages of DR. It correctly classified a large proportion of "No DR" and "Proliferative DR" cases with high specificity and sensitivity at the extremes of the disease spectrum. Mild confusion was, however, observed between "Moderate" and "Severe" classes because of the morphological resemblance between lesions at those grades, e.g., tightly packed microaneurysms and exudates.

Final Output

The Figure: 3, Figure: 4, Figure: 5 shows the web interface of the proposed model and its real time implementation. This showcases the model's ability for accurate detecting the severity of Diabetic Retinopathy across all five stages NO DR, Mild Stage DR, Moderate Stage DR, Severe stage DR and Proliferative stages on real time OCT scan fundus images.

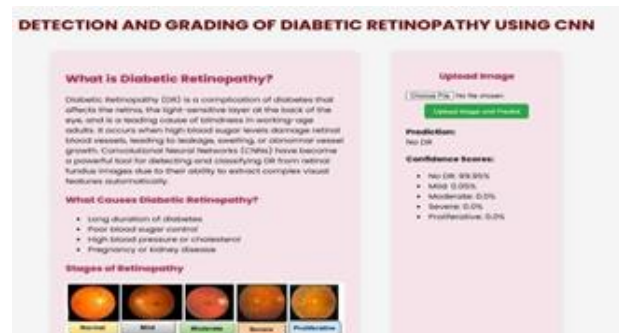


Figure: 3 Screenshot of real time implementation showing NO DR Stage



Figure: 4 Screenshot of real time implementation showing MILD Stage

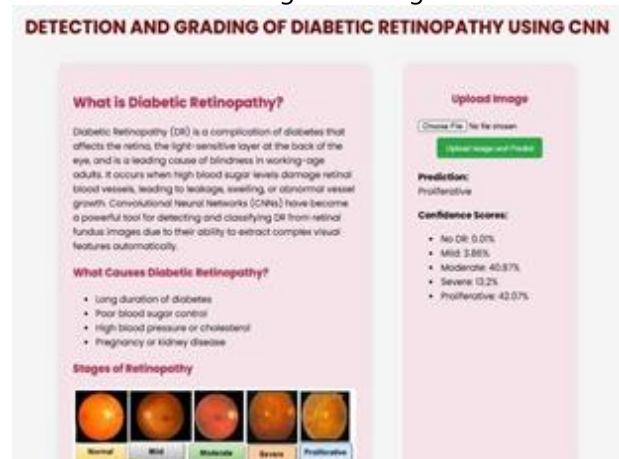


Figure: 5 Screenshot of real time implementation showing PROLIFERATIVE Stage

V. CONCLUSION

In this research, an automated deep learning approach based on the EfficientNetV2B3 model is designed for Diabetic Retinopathy (DR) detection and grading from retinal fundus images. The model is trained and validated with a five-class grading system corresponding to clinical DR severity grades: NO DR, Mild Stage DR, Moderate Stage DR, Severe Stage DR, and Proliferative Stage DR. Advanced pre-processing methods, such as contrast enhancement and image normalization, are also used to enhance feature visibility and provide consistent input quality.

The experimental outcomes have proved the model has an excellent classification accuracy, with macro-averaged precision, recall, and F1-scores over 88%, validating the model's stability in detecting both early and advanced DR stages. The EfficientNetV2B3 architecture, compared to conventional CNN models, yielded better performance based on its optimized compound scaling, light weight, and improved feature extraction capabilities. This study underscores the capability of CNN-based systems as practical diagnostic tools in the clinical environment, facilitating early detection, minimization of diagnostic workload, and ophthalmologist decision-making support. Although there are encouraging outcomes, the study also revealed limitations in the form of class imbalance and occasional misclassifications among mid-stage DR classes. A future advancement addressing these limitations using ensemble methods, attention, or lesion-aware modelling could potentially improve diagnostic accuracy. In summary, the model offered is a worthy step toward scalable and intelligent diabetic retinopathy screening, presenting an accurate, efficient, and automated early disease management and prevention of blindness solution.

Acknowledgement

We express our sincere gratitude to Dr. V. A. Kulkarni, Principal, KLS VDIT, Haliyal for their continuous support. We thank Dr. Mahendra M. Dixit, Head of Department, Electronics and Communication Engineering, KLS VDIT, Haliyal for their valuable guidance. We sincerely acknowledge our guide Prof. Nikhil Kulkarni, Assistant Professor, Dept. Electronics and Communication Engineering, KLS VDIT, Haliyal for his invaluable guidance and critical insights during the research phase.

REFERENCES

1. S. A. Jerome, "Diabetic Retinopathy Classification Using Deep Learning," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 4, pp. 399–407, 2024. [Online]. Available: [_Diabetic_Retinopathy_Classification_Using_Deep_Learning](#).
2. R. Kommaraju and M. S. Anbarasi, "Diabetic Retinopathy Detection Using Convolutional Neural Network with Residual Blocks," *Biomed. Signal Process. Control*, vol. 87, p. 105494, 2024. [Online]. [detection-of-diabetic-retinopathy-using-deep-learning](#).
3. S. V. Chilukoti, A. S. Maida, and X. Hei, "Diabetic Retinopathy Detection Using Transfer Learning From Pre-Trained Convolutional Neural Network Models," *IEEE J. Biomed. Health Inform.*, vol. 20, pp. 1–10, 2022. [Online]. [diabetic-retinopathy-detection-using-transfer-learning-from-pre-trained-convolutional-neural-network-models](#).
4. E. AbdelMaksoud, S. Barakat, and M. Elmogy, "A Computer-Aided Diagnosis System for Detecting Various Diabetic Retinopathy Grades Based on a Hybrid Deep Learning Technique," *Med. Biol. Eng. Comput.*, vol. 60, pp. 2015–2038, 2022. [Online].
5. H. Zhao, J. Jia, and V. Koltun, "A Deep Learning System for Predicting Time to Progression of Diabetic Retinopathy," *Nat. Med.*, vol. 30, pp. 1–10, 2024. [Online]

6. X. Xia, K. Zhan, Y. Fang, W. Jiang, and F. Shen, "Lesion-aware network for diabetic retinopathy diagnosis," arXiv preprint arXiv:2408.07264, 2024. [Online]. Available: .
7. M. R. Shoaib et al., "Deep Learning Innovations in Diagnosing Diabetic Retinopathy: The Potential of Transfer Learning and the DiaCNN Model," arXiv preprint arXiv:2401.13990, 2024. [Online].
8. Al-Kamachy, R. Hassanpour, and R. Choupani, "Classification of Diabetic Retinopathy using Pre- Trained Deep Learning Models," arXiv preprint arXiv:2403.19905, 2024. [Online]. Available:
9. M. S. H. and A. A. Bosale, "Detection and Classification of Diabetic Retinopathy using Deep Learning Algorithms for Segmentation to Facilitate Referral Recommendation for Test and Treatment Prediction," arXiv preprint arXiv:2401.02759, 2024. [Online].
10. S. A. G. Sabo et al., "Detection of Diabetic Retinopathy Using Deep Learning," Cureus Journal of Medical Science, vol. 16, no. 11, p. e1795, Nov. 2024. [Online].