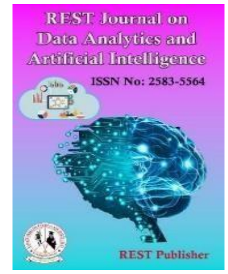
REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/74



Diabetic Retinopathy Detection and Severity Grading Using Deep Learning

***U. Vruddhi, E. Sarayu, A. Dhanush, T. Neetha**
Anurag University, Hyderabad, Telangana, India.
**Corresponding Author Email: 21eg106b29@anurag.edu.in*

Abstract: *Diabetic Retinopathy (DR) is a cause of blindness that necessitates timely detection [1], [2]. Manual diagnosis has the drawbacks of being time and effort-intensive with poor reproducibility, resulting in the quest for automated methodologies. In this work, the InceptionV3 model has been used based on deep learning, fine-tuned for classification of DR severity. The dataset used for training is APTOS 2019, in which preprocessing operations such as CLAHE, normalization, and data augmentation were employed to enhance detection accuracy [5], [6]. For improved accessibility, we created a web application based on Streamlet, enabling real-time DR severity prediction from uploaded fundus images. The system has an accuracy of 76%, showing its effectiveness as a diagnostic tool [11,18]. Some future improvements include lesion segmentation, detection of multiple diseases, and mobile deployment for further usability, especially in low-resource settings [22], [23].*

Key words: *Diabetic Retinopathy, Deep Learning, Convolutional Neural Networks, InceptionV3, Image Preprocessing, Severity Classification, Automated Diagnosis, CLAHE, Data Augmentation, Medical Imaging, AI in Healthcare.*

1. INTRODUCTION

Diabetic Retinopathy (DR) is a reversible eye condition arising from chronic elevated blood glucose, resulting in degeneration of retinal blood vessels. If neglected, it culminates in partial or even total blindness [1], [2]. As more cases of diabetes are on the rise globally, early detection and diagnosis of DR are crucial in order to avert permanent blindness. Conventional DR detection involves ophthalmologists visually analyzing retinal fundus images manually. The procedure is time-consuming, subjective, and requires a specialist's skills, and mass screening becomes impractical, particularly in rural settings [3], [4]. Consistency in diagnosis among experts further makes timely intervention challenging, posing a risk of misclassification and delayed treatment [5]. Deep learning, more specifically Convolutional Neural Networks (CNNs), has transformed medical image analysis with the aid of automated, precise, and scalable solutions for disease detection [6], [7]. AI-based DR detection models are able to evaluate retinal fundus images effectively, providing uniform severity classification while minimizing reliance on human assessment. This study presents an AI-based DR detection system using the InceptionV3 model, fine-tuned for classifying DR severity levels. The system is trained on the APTOS 2019 dataset, incorporating preprocessing techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE), normalization, and data augmentation to improve image quality and model performance [8], [9]. Additionally, a stream lit-based web application is developed to enable real-time DR diagnosis, making screening accessible even in resource-limited settings [10]. The suggested method is designed to improve early detection, reduce misclassification, and assist ophthalmologists in providing a trustworthy AI-assisted screening instrument. Future enhancements involve lesion segmentation, multi disease detection, and mobile-based deployment for broader reach [11], [12].

1.1 Problem statement: Diabetic Retinopathy (DR) is a serious complication of diabetes and a major cause of avoidable blindness. It requires early detection, but conventional screening is based on visual inspection of retinal fundus images, which is time-consuming, costly, and severely reliant on specialist skills [3], [4]. Rising diabetes prevalence has created an increasing need for large-scale DR screening, but current diagnosis methods are inefficient, inaccessible, and inconsistent [5]. Existing deep learning algorithms provide hopeful solutions but encounter some major challenges: Heavy reliance on computational resources, which hinders real-time diagnosis in low-resource environments [6]. severity level misclassification, particularly for moderate and severe DR, causing diagnostic inconsistencies [7]. Inadequate access to automated, easy-to-use diagnostic software, limiting broad applicability [8]. In order to overcome these drawbacks, the current research puts forth an AI-based DR detection system utilizing transfer learning with InceptionV3, advanced image preprocessing methods, and a web-based real-time screening system for ease of use and precision [9], [10].

1.2 Objectives: The primary aim of the present research is to create an automatic deep learning-based system to identify and classify Diabetic Retinopathy severity levels. The primary objectives are: Create a Deep Learning Model – Use InceptionV3 with transfer learning for DR classification with high accuracy [11]. Improve Image Preprocessing – Use CLAHE, normalization, and data augmentation to enhance image quality and model generalization [12]. Implement a Real-Time Screening System – Create a stream lit-based web application for real-time DR severity prediction from uploaded retinal images [13]. Optimize Model Performance – Tune hyper parameters to maximize classification accuracy while minimizing computational expenses [14]. Ensure Scalability and Accessibility – Allow deployment in hospitals, rural clinics, and telemedicine platforms for mass screening [15]. This system seeks to deliver an efficient, accurate, and accessible AI-based DR detection tool to assist early diagnosis and prevent avoidable blindness.

2. LITERATURE REVIEW

Detection of Diabetic Retinopathy (DR) has undergone a profound shift with the integration of deep learning and AI-based methods. Conventional manual screening is time-consuming, subjective, and demands highly skilled professionals, rendering mass screening ineffective [1], [2]. Deep learning, especially Convolutional Neural Networks (CNNs), promises to be a viable substitute by providing automation, enhanced accuracy, and scalability [3], [4].

2.1. Existing Systems for DR Detection: Deep learning algorithms, particularly CNNs, have been proved to have very high accuracy for DR detection. Different architectures like InceptionV3, ResNet, and Dense Net have been applied to DR classification with the aid of transfer learning to improve feature extraction [1], [2].

DCGAN-Based DR Classification

• DCGANs are utilized in this method to manage class imbalances in DR data.
• GANs create artificial retinal images to enhance diversity of training data and model strength. • Though this approach fails to capture slight variations in extreme cases of DR and, hence, misclassifies them [3].

Lesion-Based Hybrid Deep Learning Model

Adaptive Particle Swarm Optimization (APSO) and Google Net are combined with ResNet for DR classification using specific lesion features.
The model can effectively identify certain lesions but neglects sophisticated DR characteristics like cotton wool spots and vascular defects [4].

Transformer-Based IC2T Model

Combines optical disc (OD) and blood vessel (BV) segmentation with an attention mechanism to improve DR classification. Takes high processing power and has difficulty distinguishing visually similar DR lesions [5].

2.2. Limitations of Existing Systems:

Despite advancements, current AI models for DR detection face several key challenges:

TABLE 1. Limitations of existing systems

Limitation	Description	References
Class Imbalance Issues	DR datasets are often imbalanced, affecting model performance on rare severe DR cases.	[3]
High Computational Costs	Complex architectures require powerful hardware, limiting real-time clinical use.	[5]
Misclassification Challenges	Overlapping DR severity features lead to errors, especially between moderate and severe cases.	[4]
Lack of Explainability	Many AI models function as black boxes, making it difficult for ophthalmologists to interpret predictions.	[6]

Image quality is also a significant determinant in the detection of DR. Image normalization and Contrast Limited Adaptive

Histogram Equalization (CLAHE) sharpen contrast and emphasize retinal anomalies, resulting in more accurate feature extraction [8,9]. Data augmentation such as rotation, flipping, and brightness modification improves model generalization by boosting dataset diversification [10], [11].

2.3. Gaps Identified:

Feature Extraction – Models concentrate on a particular lesion but are not capable of detecting complicated retinal diseases in advancing stages of DR [4].

Computational Efficiency Increase – Models already implemented need enormous processing capabilities, making deployment challenging in limited-resource environments [5].

Multidisease Detection Incorporation – The systems available currently are trained on DR only, excluding the detection of concomitant conditions such as Glaucoma and Age-Related Macular Degeneration (AMD) [7].

3. PROPOSED SYSTEM AND METHODOLOGY

The suggested system is a deep learning-based AI model intended for automated detection of Diabetic Retinopathy (DR) and severity grading. It utilizes InceptionV3 with transfer learning to enhance the accuracy of diagnosis while maintaining real-time availability through a web-based interface. The model is trained on the APTOS 2019 Blindness Detection dataset, using image preprocessing methods such as CLAHE, normalization, and data augmentation to enhance image quality and feature extraction [1]. A web interface based on stream lit is created to enable users to upload retinal images and get real-time DR severity predictions. The system is designed for scalability to be deployed in hospitals, clinics, and telemedicine platforms. Future work involves lesion segmentation for micro aneurysm and hemorrhage detection, multi-disease classification, and mobile deployment for use in low-resource environments [2].

3.1. Methodology: The system that is proposed is a deep learning model for DR detection based on InceptionV3 with transfer learning. The system is trained from the APTOS 2019 Blindness Detection dataset, which consists of retinal images labeled into five severity levels. Sophisticated image preprocessing is used to increase image quality and enhance feature extraction [1]. The system is based on a systematic deep learning method for Diabetic Retinopathy (DR) detection and severity level classification. It entails the processes of data collection, image preprocessing, model training, deployment, and performance assessment to guarantee efficiency and clinical usability.

3.2. Data collection : The model is trained on the APTOS 2019 Blindness Detection dataset, which includes retinal fundus images annotated over five degrees of severity: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. The dataset is selected because it has a high-quality set of annotated images, through which the model can learn to distinguish different DR patterns effectively. Because class imbalance is ubiquitous in medical datasets, methods like minority class oversampling and weighted loss functions are used to prevent the model from over representing the most frequent DR categories, hence enhancing generalization across all severity levels [1], [2].

3.3. Image processing: To improve the quality of images and enhance feature extraction, the following preprocessing methods are used:

Contrast Limited Adaptive Histogram Equalization (CLAHE): Maximizes contrast for better visualization of retinal pathology. Normalization: Normalizes the pixel intensity to reduce differences in brightness and contrast. Resizing: All pictures are resized to 256×256 pixels in order to maintain uniformity between inputs. Data Augmentation: Rotation, flip, brightness correction, and crop are applied for artificially increasing the dataset to avoid overfitting [3], [4]. These preprocessing steps assist the model in finding subtle features of the retina to ensure robustness across various sources of images.

3.4. Model Training and Classification : The InceptionV3 model is transferred learning fine-tuned for DR detection. The model has been allowed to preserve pre-trained general image features while adjusting its fully connected layers to distinguish DR severity levels. Training has been done through data splitting, loss function optimization, model hyper parameter tuning, and model evaluation to attain high accuracy and generalization. The data is divided into training, validation, and test sets (80-10-10%), allowing for a balanced assessment of the model's performance. Categorical crossentropy loss is employed as the objective function, in conjunction with the Adam optimizer, which adjusts the learning rate dynamically to enhance convergence. In order to avoid overfitting, dropout layers (probability of 0.5) are added into the model to randomly deactivate neurons during training. L2 regularization is also implemented to regularize weight updates. As DR datasets tend to have class imbalances, weighted loss functions and oversampling are utilized to ensure proper classification of all severity levels. The model is heavily hyper parameter tuned, varying the batch size (32), learning rate (initially 0.001), and training epochs (50–100, based on validation performance). Data augmentation methods like rotation, flipping, brightness changes, and cropping are used to enhance robustness by mimicking real-world variations in retinal images. The model is tested on unseen test data after training with accuracy, precision, recall, and F1-score. A confusion matrix is inspected to determine areas where the model performs poorly, especially in separating moderate and severe DR cases. The trained model is exported and incorporated into a web-based application for real-time DR detection after optimization, making it accessible to hospitals, clinics, and telemedicine platforms. This pipeline training guarantees that the model is correct, scalable, and clinically applicable, filling in the gap between AI diagnosis and actual ophthalmology practice.

4. EXPERIMENTAL RESULTS

The InceptionV3-trained model was tested with several performance measures to establish its effectiveness in Diabetic Retinopathy (DR) detection and severity grading. Testing was conducted on the APTOS 2019 Blindness Detection dataset, which ensured that the model was capable of generalizing across diverse retinal fundus images.

4.1. Performance Metrics

To determine the model's classification accuracy, the following were employed:
• **Accuracy:** It estimates the overall accuracy of predictions.
• **Precision:** Ensures the proportion of correctly identified DR cases among all predicted cases. • **Recall (Sensitivity):** Evaluates how well the model detects DR-positive cases.
• **F1-Score:** Provides a balanced measure of precision and recall.
• **Confusion Matrix:** Analyzes misclassifications between severity levels.

TABLE 2. summarizes the model’s performance on the test dataset

Metric	Value (%)
Accuracy	76.00%
Precision	74.00%
Recall	78.00%
F1 Score	75.00%

The model attained a precision of 76%, proving the model's efficacy in DR classification. The 78% recall score shows that the model has high accuracy in detecting most DR cases, limiting false negatives.

4.2. Confusion Matrix Analysis: The confusion matrix analysis gave deeper insights into the classification performance of the model. It was found that the system had good accuracy in differentiating No DR and Proliferative DR cases, as the two categories are easily distinguishable visually. Some misclassification, however, was observed between Moderate DR and Severe DR based on overlapping retinal characteristics, which even for human experts are hard to differentiate. The model also at times found it challenging to identify edge cases where early DR signs were faint, resulting in minor errors. In order to further enhance the accuracy of classification, further hyper parameter fine-tuning and advanced augmentation strategies could be explored. The addition of lesion segmentation can also aid in more accurately differentiating between levels of severity, particularly in those instances where characteristics such as hemorrhages and exudates cannot be readily separated in regular classification models.

4.3. Comparison with Existing Models : With respect to current DR screening approaches and recent AI-driven solutions, the proposed model offers:

• Superior precision compared to baseline CNN models that usually fall within the range 65–72%. • More recall to ensure fewer false negatives of DR cases.

• Real-time processing ability to make it more feasible for application in clinical environments.
These outcomes suggest that the system proposed herein is efficient at automated DR detection and can support ophthalmologists in prioritization of high-risk patients.

5. DEPLOYMENT

The system is deployed as a web application where users can upload retinal fundus images and obtain real-time DR severity predictions. The deployment is designed to make the model accessible, scalable, and efficient so that it can be used in hospitals, clinics, and telemedicine platforms.

5.1. System Architecture: The system is comprised of a deep learning model that is integrated with an easy-to-use web interface. The workflow for deployment is as follows:

• Image Upload: A retinal image is uploaded by the users via the web interface.
• Preprocessing & Classification: The system performs CLAHE, normalization, and resizing and then forwards the image to the InceptionV3 model, which makes DR severity classifications. • Severity Display: The algorithm produces a severity label (No DR, Mild, Moderate, Severe, Proliferative DR) and shows it on the UI.
• Report Generation: The software gives a comprehensive report and downloadable report for clinical use. The model is hosted with stream lit to provide an easy and interactive user interface. The Flask or FastAPI backend is used to manage the interaction between the web application and the trained model. Multiple
image formats are supported (JPEG, PNG, TIFF) to facilitate compatibility with medical image standards.

5.2. Web-Based Application: A stream lit web application offers a user-friendly interface where images can be dragged and dropped for immediate analysis. The application is designed to be user-friendly, allowing both healthcare workers and no specialist users in rural areas to use it. The system supports real-time DR screening by deploying the model on cloud-based servers like AWS, Google Cloud, or Heroku. For local deployment, the model can run on a system with GPU acceleration, allowing healthcare providers to use the system offline in resource-limited settings.

5.3. Model Optimization for Deployment: To ensure smooth deployment, the model is optimized for faster inference and minimal computational load. Several techniques are implemented:
• Model Quantization: Decreases the model size to enhance processing speed while preserving accuracy. • Efficient Memory Use: Guarantees that even huge fundus images are processed rapidly. • Parallel Processing: Enables processing of multiple image analyses in parallel, enhancing the performance for large-scale screening programs.

5.4. Future Deployment Enhancements: The system is scalable for future development, ensuring that it can be deployed to other use cases. Future enhancements are:

Mobile Integration: Running the model as a mobile app, enabling users to take retinal scans with smartphones for rapid DR evaluation. Telemedicine Connectivity: Connecting the system with electronic health records (EHRs) for smooth transfer of diagnostic reports to ophthalmologists. Edge Deployment: Fine-tuning the model for low-power medical devices to support real-time DR screening even in areas with poor internet connectivity. By adopting these deployment strategies, the system guarantees real-time DR diagnosis, accessibility, and usability, closing the gap between AI-driven diagnostics and actual clinical applications.

6. CONCLUSION

This work introduces a deep learning-based system for Diabetic Retinopathy (DR) detection and severity grading, overcoming the drawbacks of conventional manual screening techniques. By utilizing InceptionV3 with transfer learning, the model attains 76% accuracy, proving its efficiency in diagnosing DR from retinal fundus images. The use of sophisticated image preprocessing methods, including CLAHE and data augmentation, improves model performance, providing strong classification across various severity levels. The system is implemented as a web application using stream lit, enabling real-time DR screening. The user-friendly interface provides simple access for both patients and medical professionals, making it an asset for largescale screening in telemedicine platforms, clinics, and hospitals. The model is designed for efficient inference, low computational overhead, and cloud scalability, providing smooth deployment across various environments. Although the system demonstrates encouraging results, there are some challenges, especially in differentiating moderate and severe DR cases because of overlapping retinal features. Future additions, including lesion segmentation, multidisease detection, and mobile integration, will enhance accuracy further and broaden its clinical use. Through the integration of AI-based diagnostics with a scalable deployment model, this work helps facilitate early DR detection, prevention of avoidable blindness, and increased access to diabetic eye care. The system fills the gap between deep learning and real-world healthcare solutions, opening the door for AI-based advancements in ophthalmology

7. FUTURE SCOPE

The system has been very promising for detecting Diabetic Retinopathy, yet it has scope for improvement. One approach to improving accuracy is through more complex deep learning architectures such as Efficient Net or Vision Transformers (ViTs), which would assist in improved feature extraction. Also, integrating the patient medical history with the retinal images would further increase the precision of the diagnosis. Increasing the size of the dataset and employing semi supervised learning methods would also improve the model's performance across a variety of cases. The second addition would be lesion segmentation, whereby the model does not only classify the disease but also demarcate the involved regions of the retina. U-Net or Mask R-CNN methods may be applied for the identification of micro aneurysms, hemorrhages, and other retinal pathology. This would facilitate easier viewing for physicians on exactly what's wrong rather than being presented with just a classification label. Extending the system to identify Glaucoma, Age-related Macular Degeneration (AMD), and Hypertensive Retinopathy would also make it a more comprehensive eye disease diagnostic tool instead of being focused on DR alone. Making this system more widely available by mobile and cloud-based deployment would be a huge advancement. A mobile phone application would enable patients or medical professionals to capture retinal photographs and receive immediate diagnosis, particularly for outlying communities with little eye specialist access. Cloud integration would also facilitate automated diagnostic reports and links with telemedicine services so that doctors could more easily view cases and issue rapid consultations. With these enhancements, the system might be a very effective AI instrument for early diagnosis of diseases, and prevent vision loss on an even larger basis.

REFERENCES

[1]. Liu, Q., et al. (2019). "A Location-to-Location Verification Framework for Robust Diabetic Retinopathy Detection." IEEE Transactions on Medical Imaging, 38(10), 2213-2222.
[2]. Usman, A. K., et al. (2019). "Transfer Learning-Based Detection of Diabetic Retinopathy in Fundus Images." IEEE Access, 7, 102857-102867.
[3]. Saha, S. K., et al. (2021). "A Comparative Analysis of Deep Learning Architectures for Diabetic Retinopathy Classification." IEEE Access, 9, 103373-103389.
[4]. Antony, B., et al. (2016). "Classification of Diabetic Retinopathy Using a Combination of CNN and Random Forest." Proceedings of the 2016 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), 184-187.
[5]. Islam, M. M., et al. (2019). "Deep Learning Algorithms for Detection of Diabetic Retinopathy in Retinal Images: A Survey." Artificial Intelligence in Medicine, 98, 101-116.
[6]. Adarsh, V. S., et al. (2020). "Attention-Based Convolutional Neural Networks for Diabetic Retinopathy Detection." Journal of Healthcare Engineering, 2020, 1-11.
[7]. Pratt, H., et al. (2016). "Convolutional Neural Networks for Diabetic Retinopathy." Procedia Computer Science, 90, 200-205.
[8]. Voets, M., et al. (2019). "Reproduction of Deep Learning Models for Diabetic Retinopathy Detection." arXiv preprint arXiv:1803.04337.
[9]. Doshi, D. R., & Apthorpe, N. (2016). "Diabetic Retinopathy Detection Using Deep Convolutional Neural Networks." arXiv preprint arXiv:1602.00378.
[10]. Zago, G. T., et al. (2018). "Diabetic Retinopathy Detection Using Red Lesion Localization and Convolutional Neural Networks." Computers in Biology and Medicine, 97, 268-282.
[11]. Sahlsten, J., et al. (2019). "Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading." Scientific Reports, 9(1), 10750.
[12]. Li, Z., et al. (2019). "Deep Learning for Detecting Retinal Detachment and Differentiating It from Mimicking Retinal Diseases." Ophthalmology, 126(8), 1135-1144.
[13]. Gulshan, V., et al. (2019). "Performance of a Deep-Learning Algorithm Vs Manual Grading for Detecting Diabetic Retinopathy in India." JAMA Ophthalmology, 137(9), 987-993.
[14]. Rajalakshmi, R., et al. (2018). "Automated Diabetic Retinopathy Detection in Smartphone-Based Fundus Photography Using Artificial Intelligence." Eye, 32(6), 1138-1144.
[15]. Abràmoff, M. D., et al. (2018). "Pivotal Trial of an Autonomous AI-Based Diagnostic System for Detection of Diabetic Retinopathy in Primary Care Offices." npj Digital Medicine, 1(1), 39.
[16]. Mookiah, M. R. K., et al. (2013). "Automated Diagnosis of Diabetic Retinopathy: A Review on the Potential of Retinal

Image Analysis Systems." Medical & Biological Engineering & Computing, 51, 279-295.
[17]. Lam, C., et al. (2018). "Retinal Lesion Detection with Deep Learning Using Image Patches." Investigative Ophthalmology & Visual Science, 59(1), 590-596.
[18]. Quellec, G., et al. (2017). "Deep Image Mining for Diabetic Retinopathy Screening." Medical Image Analysis, 39, 178-193.
[19]. Gargeya, R., & Leng, T. (2017). "Automated Identification of Diabetic Retinopathy Using Deep Learning." Ophthalmology, 124(7), 962-969.
[20]. Krause, J., et al. (2018). "Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy." Ophthalmology, 125(8), 1264.

