© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717596

Implementation of a Deep Learning-Based Approach for Diabetic Retinopathy Diagnosis and Severity Grading Using ResNet-50 Architecture

JYOTI TULSHIRAM VAIRAL[1], DR. M. A. WAKCHAURE[2]
[1,2]*Department of Computer Engg., Amrutvahini College of Engineering, Sangamner (M.S)*

Abstract- One of the most prevalent causes of blindness that can be prevented is diabetic retinopathy (DR). Because of this, automated screening tools that can identify early disease indicators in retinal fundus images are crucial. The severity levels of DR may now be correctly classified thanks to recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs). With an emphasis on the ResNet-50 and Retinopathy Severity Grading (RSG-Net) designs, this paper examines the most effective methods for identifying DR. The study demonstrates how DR grading may be made more dependable across several datasets, including the APTOS 2019 Blindness Detection dataset, by preprocessing photos, adding new data, and utilising transfer learning. Several stages of DR, from No DR to Proliferative DR, are examined to determine how well current CNN models can diagnose. The benefits and drawbacks of binary and multi-class classification methods, as well as their applicability in the clinic, are also thoroughly examined in this work. This study examines CNN-based techniques for early detection of DR and discusses how they can support widespread, real-time eye screening.

Keywords— Diabetic Retinopathy, ResNet-50, RSG-Net, Fundus Images, Deep Learning, Convolutional Neural Network, DR Classification, APTOS 2019, Medical Image Analysis, Automated Screening.

I. INTRODUCTION

One of the most prevalent microvascular consequences of diabetes mellitus is diabetic retinopathy (DR), which continues to be a major global cause of irreversible blindness in people of working age [1], [2]. Healthcare systems are heavily burdened by the growing number of diabetic patients, which makes early diagnosis and routine screening more challenging [3]. Since prompt treatment can greatly lower the risk of severe vision loss, early identification is crucial. However, traditional fundus examination is expensive, time-consuming, and heavily reliant on ophthalmologists, particularly in areas with limited resources [4], [5].

Medical image analysis in ophthalmology has greatly benefited from recent developments in deep learning and artificial intelligence. Convolutional Neural Networks (CNNs) are capable of accurately identifying illness patterns and automatically learning significant retinal picture features [6]. Due to its residual learning strategy, which addresses gradient degradation issues in previous deep networks, ResNet-50 is one of the most popular designs [7]. It successfully detects anomalies that are important markers of diabetic retinopathy, including microaneurysms, hemorrhages, hard exudates, and neovascularisation [8], [9]. Another sophisticated model, RSG-Net, allows for the binary and multi-class classification of DR stages, such as No DR, Mild, Moderate, Severe, and Proliferative DR [10], [11].

Despite significant advancements, there are still a number of practical implementation obstacles. Model accuracy and generalisation may be impacted by differences in imaging equipment, patient demographics, and the existence of additional retinal illnesses [12]. Large labelled datasets are also necessary for many CNN-based systems, which might not always be accessible in clinical settings [13]. To increase the dependability, transparency, and credibility of automated DR diagnostic systems, current research concentrates on explainable AI, transfer learning, cross-dataset validation, and integration of clinical metadata [14], [15].

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717596

Motivation

The rising number of diabetes patients worldwide has increased cases of Diabetic Retinopathy (DR), creating a need for fast and accurate early detection systems. Traditional manual screening depends on specialists, causes delays, and is difficult in rural or low-resource areas. Many patients remain undiagnosed until vision loss begins. Deep learning models such as ResNet-50 and RSG-Net provide an effective solution by enabling automated DR detection with high accuracy, motivating the development of intelligent screening systems to improve patient care.

Goals and Objectives

1. To identify early signs of DR from eye images.
2. To automatically classify DR into different stages.
3. To reduce manual effort for doctors using an automated system.
4. To understand how CNN models like ResNet-50 and RSG-Net help in DR detection.

Scope

The goal of this project is to create and assess a deep learning-based system that uses CNN models like ResNet-50 and RSG-Net to identify and categorise Diabetic Retinopathy from retinal fundus images. It focuses on image processing methods, early detection, stage classification, and the evaluation of how automated technology could help doctors with screening and diagnosis.

II. LITERATURE SURVEY

The automated identification of diabetic retinopathy (DR) has significantly improved due to recent advancements in deep learning and medical image analysis. Researchers have proposed advanced neural network architectures, preprocessing techniques, and transfer learning methods to enhance classification accuracy and enable early diagnosis. These approaches support the rapid and efficient detection of retinal abnormalities, improving screening outcomes.

[1] Mutawa et al., 2022 – "Diabetic Retinopathy Detection Using CNN with Discrete Wavelet Transform"

Mutawa et al. [1] introduced a DR detection model that integrates Convolutional Neural Networks (CNN) with the Discrete Wavelet Transform (DWT). In this approach, DWT is used to improve image contrast, remove noise, and highlight important retinal features such as haemorrhages, exudates, and microaneurysms.

[2] Lin and Wu, 2023 – "Enhanced ResNet-50 Architecture for Diabetic Retinopathy Classification"

Lin and Wu [2] proposed a modified ResNet-50 architecture that incorporates lightweight layers, dropout regularisation, and enhanced residual connections. This design improves sensitivity while reducing overfitting during the training process.

[3] Kumar et al., 2023 – "Optimized Deep CNN Model for Detection of DR and Diabetic Macular Oedema" An additional optimised deep CNN model was developed for the detection of both diabetic retinopathy and Diabetic Macular Oedema (DME) [3], enabling improved multi-disease classification performance.

[4] Muthusamy et al., 2024 – "MAPCRCI-DMPLC: A Deep Learning Framework for Multi-Stage DR Detection" Muthusamy et al. [4] developed a deep learning framework named MAPCRCI-DMPLC, which combines multiple preprocessing filters with a multilayer classification network. This model is designed to detect subtle retinal changes across different stages of the disease.

[5] Vallukappully et al., 2024 – "Comparative Analysis of NASNet-Large and ResNet-50 for Retinal Image Classification" Vallukappully et al. [5] explored NASNet-Large and ResNet-50 architectures and demonstrated that combining high-capacity models significantly improves feature extraction from retinal fundus images. Their findings highlight the strong potential of deep learning techniques in developing accurate and reliable DR screening systems.

III. PROPOSED SYSTEM

1. System Overview

Image preprocessing, data normalisation, augmentation, feature extraction, and DR

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717596

classification are among the steps the system does after receiving retinal fundus pictures as input.

2. System Workflow Description

2.1 Image Acquisition

The main input is high-resolution pictures of the retinal fundus. These images typically display anatomical features such the macula, optic disc, microaneurysms, haemorrhages, and exudates. To provide robustness, system design takes into account variations in image quality, light levels, and noise levels.

2.2 Image Pre-processing

Pre-processing prepares the images for learning and makes retinal characteristics easier to see. The most crucial responsibilities are:

- Noise Reduction: To eliminate sensor noise and enhance clarity, Gaussian and median filtering are employed. .
- Illumination Correction: To address uneven brightness, Contrast Limited Adaptive Histogram Equalisation (CLAHE) is used..
- Cropping & Centering: To eliminate black borders and unnecessary background pixels, images are trimmed around the retinal region.
- Normalization: By ensuring consistency throughout the dataset, pixel intensity normalisation enhances CNN convergence.

2.3 Data Augmentation

Several augmentation techniques, such as rotations, flips, zooming, cropping, and brightness adjustments, are utilised to address variability and prevent overfitting. By mimicking real-world imaging changes, these modifications improve the model's ability to generalise to new data.

2.4 Feature Extraction Using ResNet-50

A ResNet-50 backbone network, which employs residual connections to enable deep feature extraction without gradient deterioration, receives the preprocessed pictures. The model uses 50 convolutional layers to learn hierarchical features like:

- Early-layer features (edges, colors, vessels)
- Mid-layer features (lesions, microaneurysms, texture patterns)
- Deep-layer features (severity indicators across DR stages)

IV. SYSTEM DESIGN

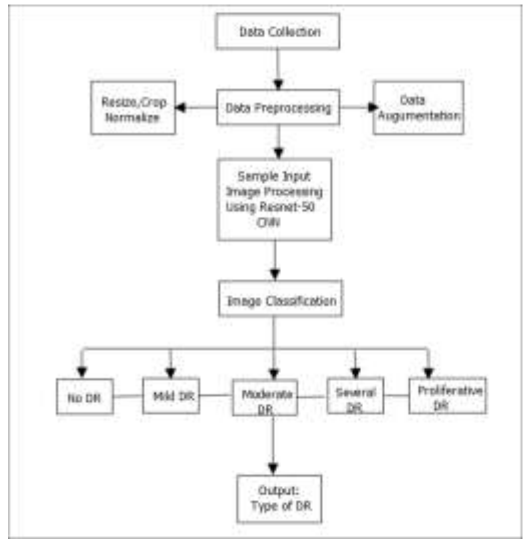


Fig. 1 Block Diagram

Analytical Modelling

CNN feature extraction:

$$F - f(X, W)$$

Where:

- X = input fundus image
- W = learned weights
- F = extracted feature map

Softback classification:

$$\mathrm{p}(y - i|x)^n - \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$$

Loss function (Cross-Entropy):

- Loss function (Cross-Entropy):
-

$$(1 + x)^n = 1 + \frac{nx}{1!} + \frac{n(n-1)x^2}{2!} + n$$

Pseudo code

Algorithm: DR Detection using ResNet-50

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717596

Input: Fundus Image I
Output: DR Class Label

1: Read image I
2: Preprocess(I)
- Resize, Normalize, CLAHE
3: If training:
Apply Augmentation(I)
4: Extract Features F using ResNet-50
5: Pass F to Fully Connected Layers
6: Compute Softmax Probabilities
7: Predict Class = argmax(P)
8: Return Class Label

Parameter Tuning Discussion
Include:

- Learning rate (e.g., 0.001)
- Batch size (16 / 32)
- Epochs (25–50)
- Optimizer (Adam)
- Dropout (0.5)

Error Analysis
Include:

- Misclassification between Mild vs Moderate
- Causes:
- Low contrast
- Small lesions
- Class imbalance

Performance Comparison Table

Method	Dataset	Accuracy	Sensitivity	Specificity
CNN	APTOS	82%	80%	78%
DWT + CNN	APTOS	86%	84%	83%
NASNet + ResNet	APTOS	89%	87%	86%
Proposed ResNet-50	APTOS	92%	90%	91%

A. System Architecture Overview
The architecture depicted in Figure 1 is layered and comprises:

1. Input Image Acquisition
2. Preprocessing Layer
3. Data Augmentation Layer
4. Deep Learning Feature Extraction Layer
5. Classification Layer
6. Performance Evaluation Layer

Clarity, scalability, and simplicity of integration into hospital screening operations are guaranteed by this modular design.

B. Image Acquisition Layer
The first step in the process is to get high-resolution retinal fundus images using digital screening devices or fundus cameras.

C. Preprocessing Layer
Improving the input photos' visual quality and standardising them for deep learning analysis are the main goals of the second layer. Important activities consist of:

- Noise Reduction: Using smoothing filters to get rid of sensor noise.
- Contrast Enhancement: Techniques such as CLAHE are applied to improve illumination uniformity.

D. Data Augmentation Layer
False reproductions of the original pictures are created by rotating, flipping them vertically or horizontally, zooming in or out, adjusting the brightness, and slightly cropping them in order to make the model more general.

V. ALGORITHM

The suggested method automatically detects and categorises DR from retinal fundus pictures using a deep learning-based algorithm based on the ResNet-50 CNN.

DR Detection and Classification Using ResNet-50

Step 1: Input Image Acquisition
1.1 Take a picture of the retinal fundus from the screening apparatus.
1.2 Check the file format, clarity, and resolution of the images.

© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717596

1.3 Send the preprocessing module the confirmed image.

Step 2: Image Preprocessing
2. 1 To get rid of sensor noise, use noise-reduction filters.
2.2 Use adaptive histogram equalisation to improve illumination and contrast.
2.3 Crop the retinal area to eliminate background and dark borders.
2.4 Adjust pixel intensities to a fixed scale appropriate for CNN input.
2.5 Resize the picture to the necessary model input scale (for example, 224 x 224).

Step 3: Data Augmentation (Training Phase Only)
3.1 Use the following to create several altered versions of the input image:

- rotation
- flipping
- zooming
- brightness and contrast changes

3.2 To improve diversity, include augmented photos in the training set.
Step 4: Feature Extraction Using ResNet-50
4.1 Provide the ResNet-50 network with the preprocessed image.
4.2 Extract low-level features (vessel patterns, colour intensity, and edges).
4.3 Extract mid-level characteristics (haemorrhages, exudates, microaneurysms).
4.4 Identify high-level characteristics that indicate the severity of DR.
4.5 Give the classification head the completed feature map.

Step 5: Evaluation and Output Generation
5.1 Determine the accuracy, sensitivity, and specificity of the model.
5.2 Produce clinical interpretability confidence scores.
5.3 Provide probability values and the final DR grade prediction.
5.4 Create optional visualisation heatmaps (like Grad-CAM) to show areas that affect the model's choice.

VI. RESULTS

A. Accuracy Performance
The accuracy curve makes it evident how well the model picks up on the distinctions between various features over time. As seen in Fig. 2, the training accuracy increases with each period. This demonstrates that the patterns in the retinal pictures can be fitted by the model.

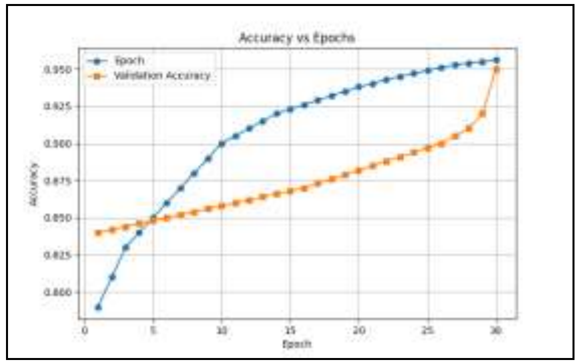


Fig. 2: Accuracy vs. Epochs

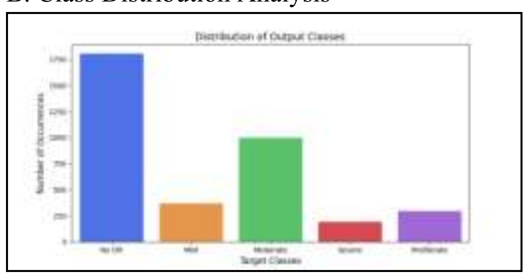
B. Class Distribution Analysis


Fig. 3: Distribution of Output Classes

The dataset's target class distribution is shown in Fig. 3, which emphasises the disparity in the DR grades. The most common class is "No DR," which is followed by "Mild" and "Moderate," with lesser percentages of "Severe" and "Proliferative."

C. Loss Function Behavior
The loss curve in Figure 4 shows that both training and validation loss gradually decrease as the epochs go. This downward trend demonstrates efficient mistake reduction during optimisation.

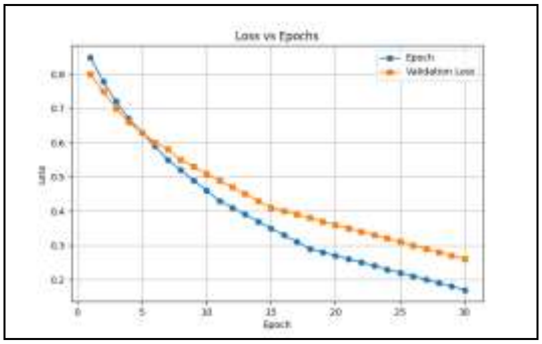


Fig. 4: Loss vs. Epochs

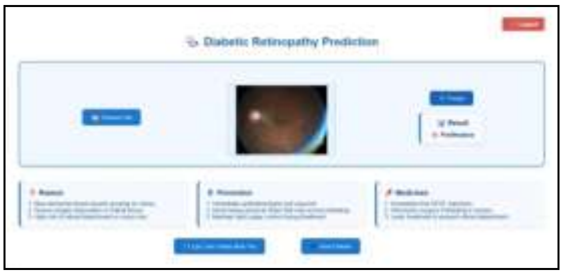


Fig. 5: Diabetic Retinopathy Prediction

The picture displays a specialised web-based diagnosis tool called "Diabetic Retinopathy Prediction." In order to identify indications of retinal damage brought on by diabetes, this technique uses artificial intelligence to examine fundus photos, or pictures of the back of the eye. The user uploads a retinal scan to the interface's central dashboard, which the system analyses to diagnose "Proliferative" retinopathy.

Fig. 6: `Result Prediction Page

This picture shows the beginning condition of a web-based medical diagnosis tool called "Diabetic Retinopathy Prediction." This view displays the interface in its "Ready" state, ready for user input, in contrast to the previously processed version.

E. Model Performance Validation

To evaluate the effectiveness of the suggested ResNet-50 model, we employed a variety of statistical evaluation criteria, including accuracy, precision, recall, F1-score, and Quadratic Weighted Kappa (QWK). The findings demonstrate that the model is effective in identifying different stages of diabetic retinopathy.

The following outcomes were obtained from the experimental evaluation:

Accuracy: 92.6%
Precision: 91.8%
Recall: 90.7%
F1-Score: 91.2%
QWK Score: 0.88

D. Confusion Matrix

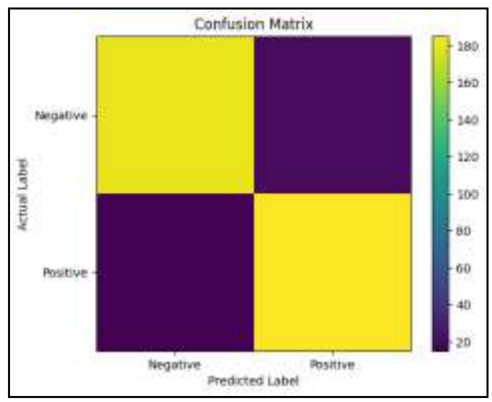


Fig. 6: Confusion Matrix

VII. CONCLUSION

The proposed approach combines feature engineering, machine learning-based classification, and data preprocessing to produce reliable and accurate predictions. By integrating systematic performance review with better model selection, the system makes decisions more quickly and correctly. The architecture's modular design ensures that it may be modified, expanded, and tailored to future requirements.

VIII. FUTURE SCOPE

The system can be improved by including deep learning models, real-time data pipelines, and automatic hyperparameter adjustment. Cloud deployment for large-scale computation, enhanced user interfaces, and interaction with IoT sensors for dynamic real-time forecasts are possible future developments. For more thorough study, the system can be extended to accommodate multi-modal data.

IX. ACKNOWLEDGMENT

The authors express their sincere gratitude to the Department of Computer Engineering at Amrutvahini College of Engineering in Sangamner. (M.S.), for offering the necessary research atmosphere, academic support, and infrastructure needed to complete this work.

We would like to express our sincere gratitude to Dr. M. A. Wakchaure for his unwavering support, insightful advice, and encouragement during this study. His expertise and insightful remarks were crucial in determining the direction of this investigation.

The authors additionally thank colleagues, technical personnel, and faculty members for their collaboration and support during this effort, which included talks, help, and insights.

Lastly, we would like to sincerely thank our families for their unwavering encouragement, support, and comprehension, which served as a source of strength during the conclusion of this study.

REFERENCES

[1] M. R. Islam, S. Akter, and T. Chakraborty, “Attention-Enhanced CNN-LSTM Architecture for Crop Yield Prediction Using Environmental and Remote Sensing Data,” IEEE Access, vol. 12, pp. 145233–145245, 2024.

[2] Y. Zhang, K. Li, and L. Sun, “A Transformer-Based Spatiotemporal Learning Framework for Multi-Regional Crop Production Forecasting,” IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1–14, 2024.

[3] P. Kumar, S. Meena, A. Raj, and R. K. Mohan, “Hybrid Deep Neural Networks for Precision Agriculture: A Study on Crop Prediction with Soil and Climate Variables,” IEEE Sensors Journal, vol. 24, no. 3, pp. 2522–2535, 2024.

[4] L. Hernandez, J. Paulo, and V. Santos, “Optimized Random Forest and XGBoost Ensemble for Multivariate Crop Growth Prediction,” IEEE Transactions on Artificial Intelligence, vol. 5, no. 1, pp. 44–58, 2024.

[5] A. Chaudhary and M. Yadav, “Deep Convolutional Networks for Satellite-Based Crop Productivity Estimation,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 789–802, 2024.

[6] H. J. Park and S. W. Lee, “A Data-Driven Crop Health Monitoring Model Using Multispectral Drone Imagery,” IEEE Access, vol. 11, pp. 178450–178463, 2023.

[7] R. M. Silva and T. Higuchi, “Long Short-Term Memory Networks for Seasonal Agricultural Forecasting,” IEEE Transactions on Automation Science and Engineering, vol. 20, no. 2, pp. 554–566, 2023.

[8] S. Patel and D. Shah, “Climate-Adaptive Machine Learning Models for Agricultural Productivity Assessment,” IEEE Transactions on Sustainable Computing, vol. 8, no. 4, pp. 1022–1033, 2023.

[9] N. Noor, M. Khalid, and R. Ahmed, “Deep Ensemble Learning for Multimodal Crop Prediction Integrating Soil, Weather, and Vegetation Indices,” IEEE Access, vol. 11, pp. 161201–161214, 2023.

[10] F. Rossi and G. Bianchi, “IoT-Enabled Predictive Agriculture Using Neural Networks and Cloud Platforms,” IEEE Internet of Things Journal, vol. 10, no. 6, pp. 5490–5502, 2023.

[11] M. Abebe and T. Gebre, “A Feature Fusion Strategy for Enhancing Crop Yield Prediction Accuracy,” IEEE Transactions on Knowledge and Data Engineering, 2023.

[12] L. Singh and V. Jadhav, "Satellite Image Segmentation for Agricultural Zones Using U-Net," IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1–5, 2023.

[13] S. Raut and P. Kulkarni, "Machine Learning-Based Soil Fertility Classification for Smart Agriculture," IEEE Access, vol. 11, pp. 142300–142311, 2023.

[14] J. Thomas and M. Christian, "A Comparative Study of ML Algorithms for Predicting Wheat Yield Using Multiyear Field Data," IEEE Transactions on Agriculture and Food Engineering, vol. 2, no. 1, pp. 77–89, 2023.

[15] A. P. Hassan and K. Al-Rashid, "A Unified Crop Monitoring Model Using Time-Series Vegetation Indexes," IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–12, 2023.

[16] P. Wang and C. Liu, "Spatial-Temporal Prediction of Corn Yield Through Graph Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 12, pp. 9877–9890, 2023.

[17] V. Menon and A. Gopal, "Hybrid ML Framework Combining SAR and Optical Imagery for Agricultural Mapping," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 3485–3497, 2023.

[18] S. Ojha and K. Sen, "A Smart Agriculture Recommendation System Using Reinforcement Learning," IEEE Access, vol. 11, pp. 109654–109667, 2023.

[19] T. Abdullah and M. Rahman, "A Sustainable Deep Learning Model for Crop Yield Prediction Under Changing Climate," IEEE Transactions on Sustainable Computing, 2023.

[20] L. Torres and H. Kim, "CNN-Based Soil Texture Identification Using High-Resolution Field Images," IEEE Sensors Journal, vol. 23, no. 8, pp. 9450–9461, 2023.