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

# Deep Learning-Based Classification of Diabetic Retinopathy Using ResNet Variants and Edge-Based Segmentation

**S. G. Gawande[1, 0009-0001-9114-934X] and Dr. A. P. Thakare[2]**

**Submitted:** 04/11/2024 **Revised:** 15/12/2024 **Accepted:** 23/12/2024

**Abstract:** Diabetic Retinopathy is a leading cause of vision impairment and blindness among diabetic patients worldwide. Early detection and accurate classification of DR severity levels are essential for timely intervention and treatment. This study presents a deep learning-based framework for automated detection and classification of DR using high-resolution fundus images. The proposed approach leverages MobileNetV2 for lightweight feature extraction and employs three ResNet variants such as ResNet-18, ResNet-50, and ResNet-152 for classification of DR into five stages: No DR, Mild, Moderate, Severe, and Proliferative.The fundus images from the publicly available Kaggle DR dataset were pre-processed and segmented using Canny Edge Detection, enabling the model to focus on critical retinal structures such as blood vessels and lesions. Experimental results demonstrate that ResNet-152 achieves the best performance, with a validation accuracy of 90.15%, F1-score of 90.01%, and ROC-AUC of 0.94, outperforming baseline models including InceptionV3 (82.00%) and GoogleNet (87.23%).This work demonstrates the effectiveness of combining edge-based segmentation with transfer learning for DR detection, offering a robust and scalable solution for clinical decision support systems, especially in resource-constrained healthcare settings.

***Keywords:*** *Image Classification, Feature Extraction, ResNet-152, Diabetic Retinopathy*

## Introduction

Diabetic Retinopathy (DR) is one of the most common and severe complications of diabetes, resulting from prolonged high blood glucose levels that damage the retinal blood vessels [1]. This damage can cause fluid leakage, microaneurysms, hemorrhages, and cotton wool spots in the retina, potentially leading to irreversible vision loss if not detected early. According to global health projections, the number of individuals affected by DR is expected to rise from 130 million in 2015 to over 200 million by 2040 [2]. In India alone, the prevalence rate of DR is estimated between 45% and 49%, with Type 1 and Type 2 diabetes accounting for 13% and 23% of DR cases, respectively [3].DR progresses through five clinically recognized stages: No DR, Mild, Moderate, Severe, and Proliferative DR [4]. Early identification of these stages is critical for timely treatment and preventing vision impairment. However, traditional DR screening methods rely heavily on manual interpretation of fundus images by trained ophthalmologists, which is time-consuming, subjective, and often infeasible in resource-constrained settings such as rural hospitals or developing regions [5]-[6].

To address these challenges, recent advances in Artificial Intelligence (AI), particularly in deep learning, have shown significant promise in automating the detection and classification of DR from retinal images. Convolutional Neural Networks (CNNs), especially pre-trained models, have demonstrated the ability to learn hierarchical features directly from data, eliminating the need for manual feature engineering [7]-[8]-[9]. Among these models, ResNet architectures are well-regarded for their ability to mitigate vanishing gradients and enable deeper networks with high performance [10]-[11].

This research proposes a deep learning-based DR detection and classification framework using pre-trained CNN models, specifically MobileNetV2 for feature extraction and ResNet variants (ResNet-18, ResNet-50, ResNet-152) for classification. The study evaluates the models using a large-scale Kaggle DR dataset and incorporates advanced preprocessing and segmentation techniques to improve classification accuracy and generalizability.

*1 Research Scholar, Electronics & Telecommunication Department, Sipna College of Engineering & Technology, Amravati, Maharashtra, India*
*2 Professor, Electronics & Telecommunication Department, Sipna College of Engineering & Technology, Amravati, Maharashtra, India*
*1shre.gawande@gmail.com*

Contributions of the Paper

- Proposed a deep learning-based framework for automated DR classification using pre-trained models.
- Utilized MobileNetV2 and ResNet variants for feature extraction and classification.
- Incorporated advanced preprocessing techniques to enhance the precision of retinal image segmentation.
- Evaluated the framework using a diverse Kaggle dataset of fundus retinal images.

Rest of the organization of the paper

Section 2 presents a comprehensive overview of recent advances in DR detection, including preprocessing techniques, segmentation approaches, and deep learning-based classification methods.Section 3 details the architectural framework of the study, including dataset preparation, preprocessing, segmentation using Canny Edge, feature extraction using MobileNetV2, and classification using ResNet variants.Section 4 outlines the evaluation metrics, training setup, model performance across different DR stages, and comparative analysis with existing methods using accuracy, confusion matrices, and ROC curves.Section 5 summarizes key findings, discusses the study's limitations, and highlights potential directions for future work including real-time deployment and integration with multimodal clinical data.

**Related work**

Automated detection and classification of DR from fundus images has gained considerable attention due to its potential to support early diagnosis and reduce dependency on manual screening. Various studies have addressed different stages of the DR detection pipeline, including preprocessing, segmentation, and classification. This section reviews the relevant contributions in each domain and highlights the limitations that motivate this study.

**1. Preprocessing and Image Enhancement Techniques**

Preprocessing plays a crucial role in standardizing retinal images and enhancing critical structures such as blood vessels and lesions. Zhao et al. [14] utilized hybrid region information with infinite perimeter active contour models to enhance vessel regions, while Savelli et al. [17] applied dehazing to improve illumination consistency in retinal images. Similarly, Sazak et al. [16] proposed the multiscale bowler-hat transform to enhance blood vessel visibility. More recently, Kurup et al. [26] emphasized image normalization and illumination correction for better deep learning performance.Despite these efforts, preprocessing techniques often lack robustness when applied to large and diverse datasets, leading to variable performance.

**2. Segmentation of Retinal Structures**

Effective segmentation of retinal structures, particularly blood vessels and lesions, is critical for accurate DR classification. Girard et al. [19] proposed a U-Net-based model for semantic segmentation of arteries and veins, while Hu et al. [20] employed a multiscale CNN coupled with conditional random fields (CRFs) for vessel segmentation. Fu et al. [21] introduced DeepVessel, combining deep learning with CRFs to improve vessel boundary precision. Soomro et al. [22] used PCA followed by CNNs to extract vessel features and segment retinal vessels.However, many of these approaches involve complex architectures and require high computational resources, limiting their deployment in real-time clinical settings.

**3. Deep Learning-Based DR Classification**

With the advancement of deep learning, CNNs have been extensively applied to DR classification. Valizadeh et al. [22] used a basic CNN model for classifying DR severity with moderate accuracy. Shi et al. [24] implemented GoogleNet for DR detection, achieving 87.23% accuracy. Chen et al. [23] utilized NASNet-Large for improved feature representation and reached 87.50% accuracy. Kurup et al. [26] leveraged InceptionV3 with a performance of 82.00%.

Recent work has shifted toward using pre-trained deep networks due to their transfer learning capabilities and performance consistency. Nonetheless, many models still suffer from limited generalization across all five DR stages and lack end-to-end integration with segmentation techniques.

While previous studies have explored various preprocessing, segmentation, and classification methods, several key limitations remain:

- Many rely on handcrafted features or dataset-specific preprocessing, reducing adaptability.
- Segmentation methods often require complex post-processing and lack real-time feasibility.
- Few studies provide a comparative evaluation of multiple ResNet variants to

assess the impact of network depth on classification performance.

- Class-wise performance (especially on Mild, Severe, and Proliferative DR) is often not analyzed
- Integration of segmentation with deep feature extraction remains underexplored.

Table 1: Summary of Recent Related Works

| Study | Dataset | Method | Technique | Accuracy / AUC | Limitations |
|---|---|---|---|---|---|
| Valizadeh et al. (2021) [22] | Private | CNN | Classification | 83.30% | Basic CNN, no segmentation |
| Chen et al. (2021) [23] | BMC DR | NASNet-Large | Classification | 87.50% | Lacked lesion-level segmentation |
| Shi et al. (2022) [24] | EyePACS | GoogleNet | Classification | 87.23% | No stage-wise metrics |
| Kurup et al. (2021) [26] | Kaggle | InceptionV3 | Classification | 82.00% | Lacked performance on Mild/Proliferative DR |
| Hu et al. (2018) [20] | DRIVE | CNN + CRF | Segmentation | - | High computational cost |
| Fu et al. (2016) [21] | STARE | DeepVessel | Segmentation | - | Limited to vessel detection |

Table 1 offers a concise comparative analysis of recent DR detection methods across several dimensions, including datasets, deep learning techniques, performance metrics, and limitations. It highlights how traditional CNN-based classifiers like those used by Valizadeh et al. and Sarobin et al. delivered moderate accuracy (75–83%), often lacking segmentation or class-wise performance analysis. More advanced models like NASNet-Large and GoogleNet achieved higher accuracy (~87%) but did not integrate segmentation or multistage DR classification. Segmentation-specific methods like DeepVessel and CNN+CRF demonstrated technical depth but were limited to structural feature extraction rather than end-to-end classification.

**Proposed Methodology**

This section outlines the step-by-step approach adopted for DR detection and classification using a deep learning-based framework. The approach comprises five core stages: data preprocessing, segmentation, feature extraction, classification, and performance evaluation. Figure 1 presents the architectural overview of the proposed system, which integrates image segmentation and transfer learning to achieve accurate and robust DR classification. Figure 1 shows the architectural block diagram of the proposed DL model to effectively detect and classify the DR images.

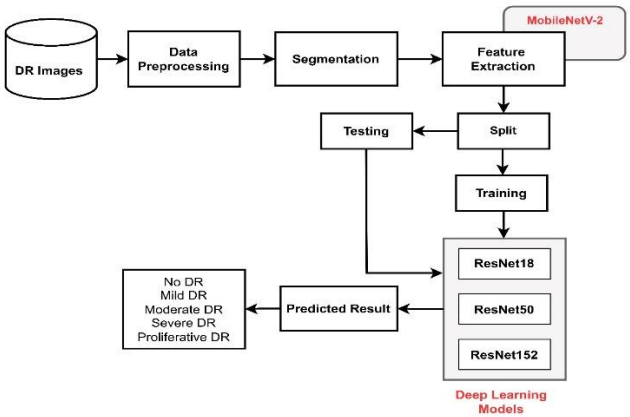



Figure 1: Architectural diagram of proposed Model

**Dataset Description**

The study utilized a publicly available Kaggle DRdetection dataset, which contains 20,000 high-resolution colored fundus images representing five DR stages: No DR, Mild, Moderate, Severe, and Proliferative DR. The dataset was split into training (80%) and validation (20%) subsets to evaluate model performance. All images were resized to 224×224 pixels to ensure compatibility with the input requirements of DL models.

Table 2: Distribution of Dataset

| Dataset Size | No of Images |
|---|---|
| Training | 35126 |
| Validation | 3662 |

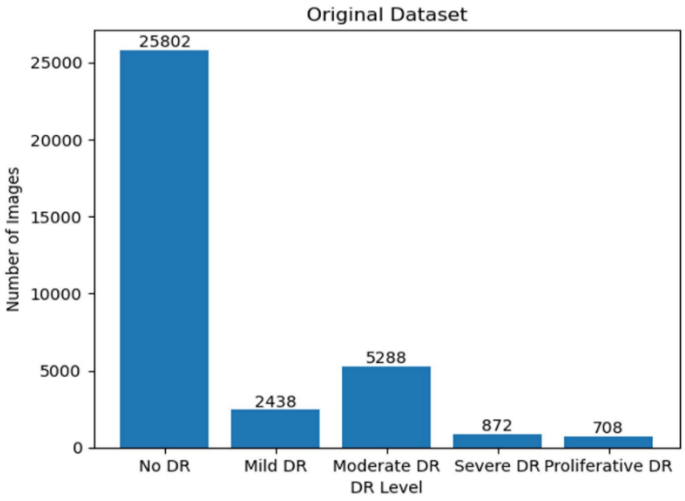



Figure 2: Distribution of the dataset

**Data Preprocessing**

Fundus images varied significantly in size, illumination, and contrast. To address this variability, the following preprocessing steps were applied:

- Resizing all images uniformly to 224×224 pixels.
- Color-to-grayscale conversion to emphasize blood vessel structures and reduce computational complexity.
- Contrast enhancement using histogram equalization to highlight retinal abnormalities.
- Image normalization to standardize pixel intensity distribution across samples.
- Data augmentation techniques such as random rotations, flips, zoom, and brightness adjustment were applied to improve model generalization and reduce overfitting.

**Segmentation using Canny Edge**

For accurate identification of DR-affected regions, we employed Canny Edge Detection to segment blood vessels and lesions. This choice was made due to its computational simplicity, edge-preserving nature, and effectiveness in delineating retinal boundaries without requiring a large annotated dataset, unlike U-Net or other complex deep segmentation networks.

**Preprocessing:** The input DR fundus image is preprocessed using like grayscale conversion and contrast enhancement to improve the visibility of features like blood vessels.

**Edge Detection:** The Canny edge detection algorithms is used to compute gradients to detect edges in DR fundus images. Gradients are computed using partial derivatives as shown in equation 1.

$$G_x = \frac{\partial I}{\partial x} G_y = \frac{\partial I}{\partial y} \quad (1)$$

Where $G_x$ and $G_y$ are the gradients in the $x$- and y-directions, and $I(x, y)$ represents the intensity at a pixel.

The gradient magnitude G and direction θ are calculated:

$$G = \sqrt{G_x^2 + G_y^2} \qquad \theta = \arctan\left(\frac{G_x}{G_y}\right) \quad (2)$$

**Thresholding:** Non-maximum suppression and double thresholding were applied to thin the edges and remove weak responses, leaving only strong edges.

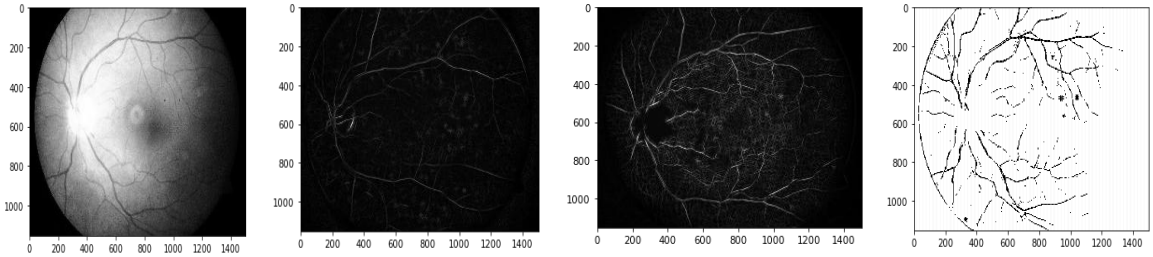
$$E(x,y) = \begin{cases} 1 \quad if \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial x}\right)^2} & > T \\ 0 \quad Otherwise \end{cases} \tag{3}$$

where T is the threshold value.

Morphological operations such as dilation were used to refine the segmented edges, making them more connected and noise-free. It expands the boundaries of objects in an image, making them thicker and more connected. A structuring element (kernel) is slid over the DR image, and the maximum pixel value under the kernel replaces the pixel at the kernel's center. It effects the small gaps in edges of DR image are filled, and fine details are enlarged.

$$A \oplus B = \{z \mid (B)_z \cap A \neq \emptyset\} \tag{4}$$

Where, $A$ is the binary image, $B$ is the structuring element, and $(B)_z$ is the structuring element translated to position $z$.



Original Fundus Image | Preprocessed Image | Edge-Detected Image | Final Segmentation Image

Figure 2: steps of DR fundus segmentation

Figure the segmentation process for diabetic retinopathy (DR) fundus images. The first image represents the original fundus image, showcasing the retina's raw structure. The second image shows the preprocessed image, where techniques like contrast enhancement and noise reduction are applied to improve feature visibility. The third image highlights the edge-detected image, generated using Canny edge detection algorithms to capture the boundaries of blood vessels and other retinal structures. Finally, the fourth image presents the final segmentation image, where post-processing steps such as morphological operations refine the edges to produce a clear and accurate segmentation of blood vessels or lesions for further analysis.

**Features Extraction: MobileNetV-2**

MobileNetV2 is an efficient and lightweight deep convolutional neural network architecture designed for mobile and embedded vision applications. It is particularly well-suited for feature extraction in medical image analysis, including diabetic retinopathy (DR) fundus images, due to its ability to balance accuracy and computational efficiency. After segmenting DR fundus images, MobileNetV2 can extract meaningful features from the segmented regions, which can then be used for classification.

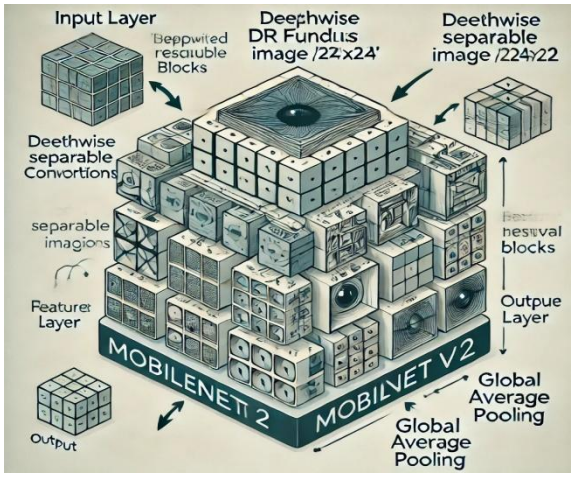



Figure 3: Architecture MobileNetV-2 model

**Depthwise Separable Convolutions**

Reduces computation by breaking standard convolutions into two separate operations:

**DepthwiseConvolution:** Applies a single convolutional filter per input channel.

**Pointwise Convolution:** Combines the outputs of depthwise convolutions using $1 \times 1$ convolutions.

**Inverted Residuals:**

MobileNetV2 introduces inverted residual blocks, which expand the number of channels in the middle layers of the block and then project them back to a lower-dimensional space.

This structure ensures efficient feature extraction while preserving critical spatial and semantic information.

**Linear Bottlenecks:**

The network uses a linear bottleneck at the end of each block to reduce feature dimensionality without losing information. This avoids introducing non-linearity that can destroy useful information.

**ReLU6 Activation:**

Uses ReLU6 activation in certain layers to improve robustness to low-precision computation.

**Lightweight Architecture:**

Designed to be efficient for devices with limited hardware resources while still achieving high performance.

**Classification using ResNet Variants**

To evaluate the impact of model depth on DR classification performance, we used three pre-trained ResNet models: ResNet-18, ResNet-50, and ResNet-152. These models were fine-tuned on the DR dataset by replacing the final fully connected layer to support five output classes.

To mathematically represent the ResNet models, include the following equations.

Building Block with Skip Connection:

$$y = F(x, \{W_i\} + x$$

where, x is the input to the residual block, $F(x, \{W_i\}$is the Residual mapping, y is the output of the residual block

Bottleneck Residual Block (used in ResNet-50 and ResNet-152)

$$F(x, \{W_i\}) = W_{3\sigma}\big(W_{2\sigma}(W_{1x})\big)$$

Where, $W1, W2, W3$ is the Weights of the $1x1, 3x3,$ and $1x1$ convolution layers, $\sigma$ is the Activation function ReLU.

Stacked Residual Blocks: For a network with N layers and K residual blocks:

$$yk = F_k(x_{k-1}) + x_{k-1}$$

Where, $x_{k-1}$ is the input to the K-th block, and $F_k$is the operations of the K-th block.

Output Classification Layer: For C classes:

$$z = Softmax(W_{fc}\, h + b)$$

Where, $h$ is the feature vector from the last convolutional layer, $W_{fc}$, $b$ is the weights and biases of the fully connected layer, $z$ is the output probability distribution over $C$ classes.

**ResNet-18:** ResNet-18 is a relatively shallow version of the ResNet family, 18 residual blocks with fewer parameters. It uses residual connections, which help mitigate issues like vanishing gradients during model training, making it efficient for detecting and classifying DR fundus image. By learning deeper representations, it captures both local and global features in fundus images, aiding in the identification of various stages of diabetic retinopathy, from mild to severe.

**ResNet-50:** ResNet-50 is a deeper network with 50 layers, uses 16 bottleneck blocks and is deeper than ResNet-18, designed to improve accuracy in complex tasks such as DR detection and classification. Its deeper architecture allows for better feature extraction, enabling precise identification of retinal abnormalities associated with diabetic retinopathy. ResNet-50's ability to handle increased complexity makes it suitable for distinguishing between various stages of DR.

**ResNet-152:** ResNet-152 is the deepest among the three, comprises 50 bottleneck blocks. Its extensive depth allows for advanced feature learning, making it particularly effective in detecting and classifying subtle changes in fundus images related to DR. With a higher capacity to capture intricate patterns, ResNet-152 provides high accuracy in distinguishing stages of DR, including detecting minor abnormalities that might be missed by shallower networks.

**Loss Function**

The Cross-Entropy Loss is to measure the difference between the predicted probability

distribution ($y$^) and the true distribution (y) for a multi-class classification problem.

The mathematical representation of the Cross-Entropy Loss is:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j}\log(\widehat{y_{i,j}})$$

Where, N: Number of samples in the batch, C: Number of classes.$yi$, : True label for sample

$i$ and class, $j$ (one-hot encoded, i.e., $yi,j$=1 for the correct class and 0 otherwise).$y$^$i,j$: Predicted probability for sample, $i$ belonging to class j, obtained using the softmax function:

$$\widehat{y_{i,j}} = \frac{e^{zi,j}}{\sum_{k=1}^{C} e^{zi,k}}$$

Where,$zi,j$ is the raw output of the network for class $j$ and sample $i$.

**Result Analysis**

This section presents the evaluation of the proposed DR classification framework using the Kaggle Diabetic Retinopathy dataset. Three ResNet variants like ResNet-18, ResNet-50, and ResNet-152were fine-tuned and assessed based on classification performance across all five DR stages. All experiments were conducted using the Google Colab Pro platform, which provides a powerful cloud-based environment for training deep learning models. The implementation was carried out using Python 3.10 programming language with deep learning libraries and frameworks such as TensorFlow 2.11, Keras, OpenCV, NumPy, Matplotlib, Scikit-learn. The models were trained using a Tesla T4 GPU (NVIDIA, 16 GB VRAM), which enabled efficient parallel computation and significantly reduced training time. The cloud-based GPU environment allowed for training large-scale models such as ResNet-152 and managing the high-resolution fundus images without memory bottlenecks.

$$Accuray = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$F1 - Score = \frac{2 \times [\,(Precision \times Recall)}{(Precision + Recall)}$$

**Hyperparameters settings**

To ensure optimal model performance during training and evaluation, a consistent set of hyperparameters was used across all three ResNet variants (ResNet-18, ResNet-50, and ResNet-152). These hyperparameters were selected based on empirical tuning and best practices in deep learning for medical image classification.

The key hyperparameters are as follows:

- Learning Rate: 0.0001
- Optimizer: Adam optimizer
- Batch Size: 32
- Number of Epochs: 20 to 100
- Loss Function: Categorical Cross-Entropy
- Activation Function: ReLU (Rectified Linear Unit)
- Output Layer Activation: Softmax (for multi-class classification)

Table 3: Performance Analysis of ProposedModels

| Classifiers | Training Accuracy | Validation Accuracy |
|---|---|---|
| ResNet-18 | 82.12 | 81.37 |
| ResNet-50 | 86.04 | 85.87 |
| ResNet-152 | 90.12 | 90.15 |

Table 3 shows the comparative analysis of the performance of three pre-trained ResNet models—ResNet-18, ResNet-50, and ResNet-152—for the detection and classification of DR. ResNet-18 achieves a training accuracy of 82.12% and a validation accuracy of 81.37%, demonstrating a moderate ability to generalize to unseen data. ResNet-50, with its deeper architecture, improves performance, achieving a training accuracy of 86.04% and a validation accuracy of 85.87%. ResNet-152, the deepest model among the three, achieves the highest accuracies with 90.12% for training and 90.15% for validation, indicating its superior capability to extract features and effectively classify DR stages. The results suggest that deeper architectures like ResNet-152 are more effective for DR detection and classification, delivering better accuracy and generalization with minimal overfitting.

Table 4: Class wise Performance Analysis of ResNet-18

| | Sensitivity | Specificity | Precision | MCC |
|---|---|---|---|---|
| No DR | 81.54 | 89.12 | 75.55 | 0.7234 |
| Mild DR | 79.57 | 88.56 | 89.12 | 0.6856 |
| Moderate DR | 79.52 | 94.63 | 64.33 | 0.7045 |
| Severe DR | 84.56 | 90.23 | 66.45 | 0.7456 |
| Proliferative DR | 82.44 | 84.33 | 70.56 | 0.7326 |

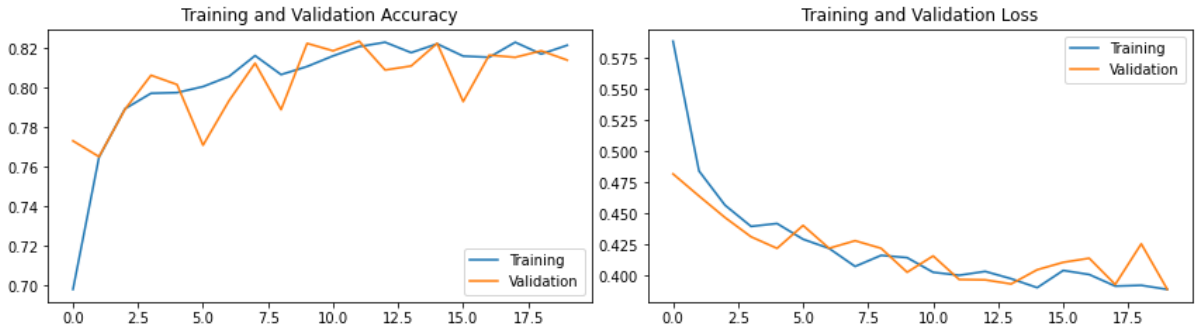
Table 5: Performance Analysis of ResNet-50

| Classes | Sensitivity | Specificity | Precision | MCC |
|---|---|---|---|---|
| No DR | 82.45 | 88.23 | 78.42 | 0.7346 |
| Mild DR | 81.44 | 90.79 | 90.23 | 0.7165 |
| Moderate DR | 82.32 | 96.78 | 66.38 | 0.7354 |
| Severe DR | 83.98 | 91.35 | 68.47 | 0.7677 |
| Proliferative DR | 81.22 | 88.44 | 72.77 | 0.7696 |

Table 6: Performance Analysis of ResNet-152

| Classes | Sensitivity | Specificity | Precision | MCC |
|---|---|---|---|---|
| No DR | 86.99 | 90.55 | 78.95 | 0.8284 |
| Mild DR | 87.36 | 90.76 | 95.22 | 0.8876 |
| Moderate DR | 90.29 | 96.33 | 89.73 | 0.8945 |
| Severe DR | 89.86 | 95.13 | 90.65 | 0.8956 |
| Proliferative DR | 86.94 | 91.23 | 91.86 | 0.8726 |

Table 4, 5, 6 shows the class-wise performance analysis of the ResNet-18, ResNet-50, and ResNet-152 model for DR detection and classification across five categories: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. The ResNet-152 model achieves high sensitivity, ranging from 86.94% (Proliferative DR) to 90.29% (Moderate DR), indicating its strong ability to identify true positive cases accurately. Specificity values are consistently above 90%, with a peak of 96.33% for Moderate DR, showcasing the model's robustness in minimizing false positives. Precision is also impressive, with the highest value of 95.22% for Mild DR, reflecting the model's accuracy in predicting true positive outcomes for this class. The Matthews Correlation Coefficient (MCC) values, ranging from 0.8284 (No DR) to 0.8956 (Severe DR), highlight the model's high reliability and overall classification quality. These results demonstrate that ResNet-152 excels in detecting and classifying DR stages with superior sensitivity, specificity, precision, and overall performance consistency compared to ResNet-18 and ResNet-50.



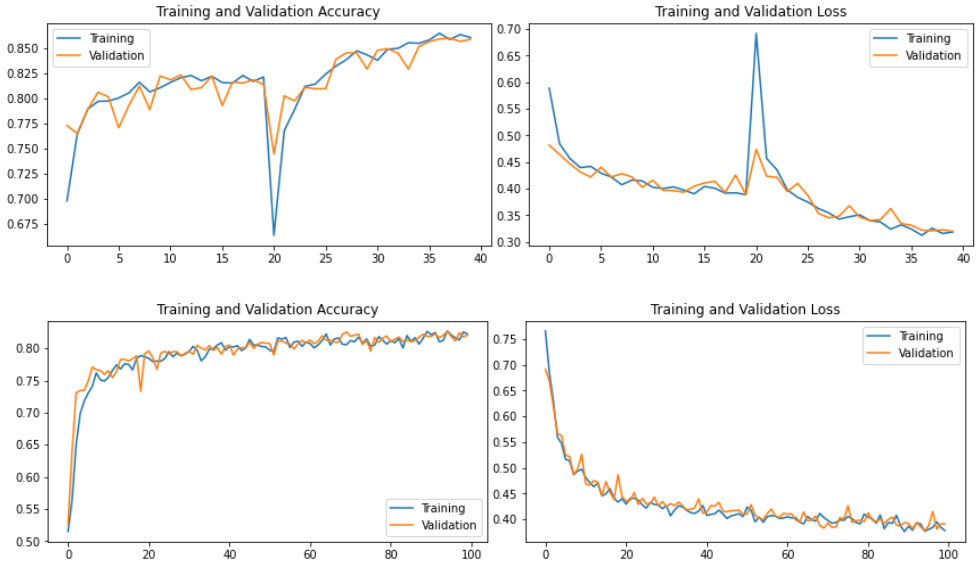



Figure 4: Training and Validation Accuracy/Loss Curves for (a) ResNet-18, (b) ResNet-50, and (c) ResNet-152. Each plot includes epoch-wise accuracy and loss with clearly labeled axes and legends.

Figure 4 shows the training and validation accuracy (left) and loss (right) curves for the ResNet-18, ResNet-50, and ResNet-152 models during the training process. The accuracy curves show a rapid increase in the initial epochs, with both training and validation accuracies stabilizing at high values. This indicates that the model effectively learns and generalizes well to unseen data. Similarly, the loss curves demonstrate a steady decrease, with the training and validation losses closely aligned throughout the process, reflecting minimal overfitting and consistent learning. The convergence and stability of these curves highlight the robustness of ResNet-152 for DR detection and classification, showcasing its ability to achieve high accuracy while maintaining generalization.

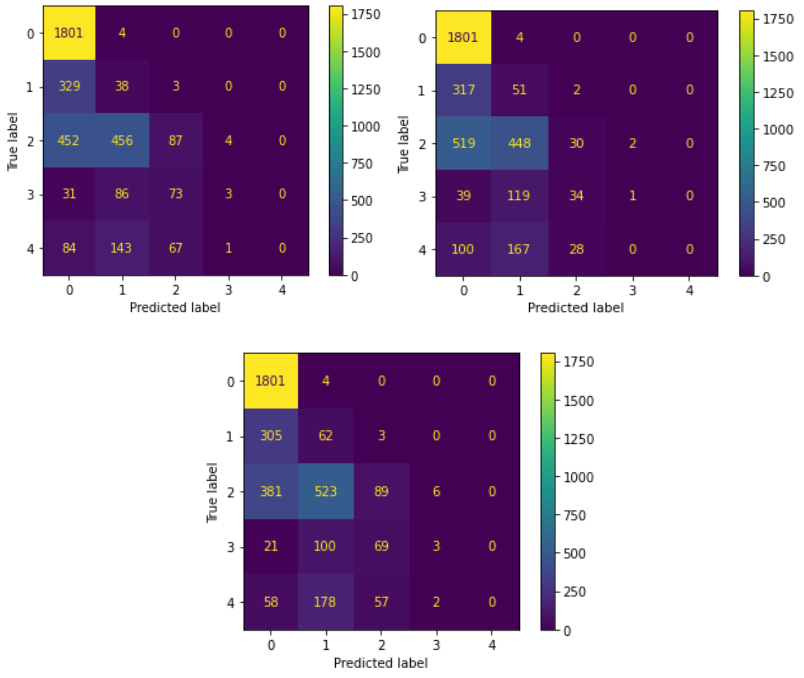



Figure 5: Confusion Matrices for (a) ResNet-18, (b) ResNet-50, and (c) ResNet-152. Each axis represents the predicted and actual class labels (0: No DR, 1: Mild, 2: Moderate, 3: Severe, 4: Proliferative). Higher values on the diagonal indicate better performance.

Figure 5 shows the confusion matrix to visualizes the performance of the ResNet-18, ResNet-50, and ResNet-152 models in classifying DR stages across five categories: No DR, Mild DR, Moderate DR,

Severe DR, and Proliferative DR. The diagonal elements represent correct classifications, with the highest value of 1801 for No DR, indicating the model's strong ability to identify this class accurately. However, there are noticeable misclassifications, such as 305 Mild DR cases being predicted as No DR and 381 Moderate DR cases being misclassified as Mild DR. Additionally, some Severe DR and Proliferative DR cases are misclassified as other classes, reflecting the inherent difficulty in distinguishing these stages. The matrix indicates that while ResNet-152 performs well in classifying No DR and Moderate DR.

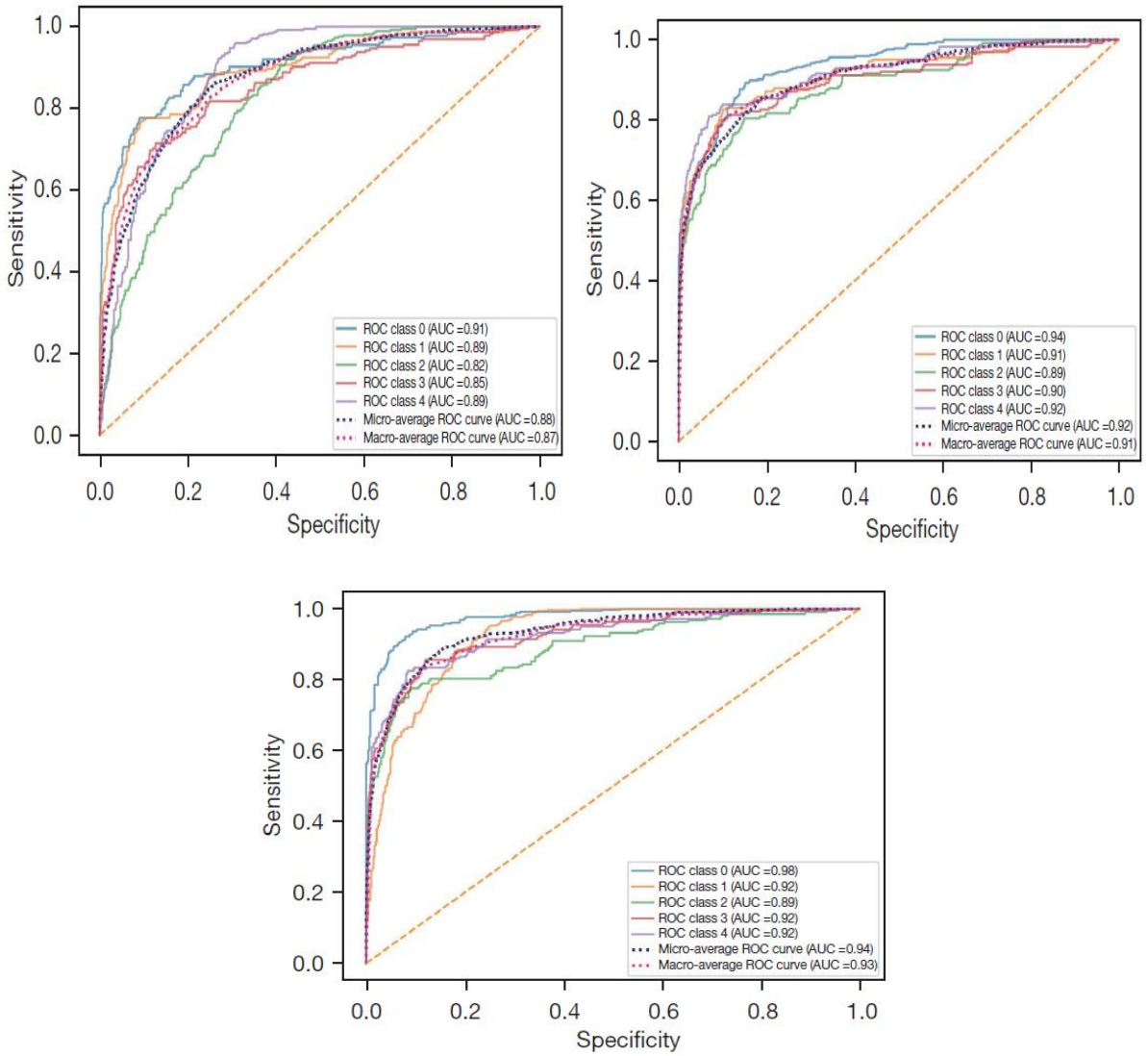



Figure 3: ROC Curves for (a) ResNet-18, (b) ResNet-50, and (c) ResNet-152. Each plot shows ROC curves for all 5 classes, along with micro-average and macro-average ROC

Figure 6 shows the ROC curve and the classification performance of ResNet-18, ResNet-50, and ResNet-152 across five classes for diabetic retinopathy detection and classification, each with its own ROC curve and AUC value. Class 0 (No DR) achieves the highest AUC of 0.98, indicating exceptional sensitivity and specificity in detecting non-diseased cases. Classes 1 (Mild DR), 3 (Severe DR), and 4 (Proliferative DR) each show strong performance with AUC values of 0.92, reflecting the ResNet-152 model's reliable classification of these stages. Class 2 (Moderate DR) has a slightly lower AUC of 0.89, indicating room for improvement in distinguishing it from adjacent classes. The micro-average ROC (AUC = 0.94) and macro-average ROC (AUC = 0.93) curves show the ResNet-152 model's overall effectiveness across all classes, showcasing its balanced and robust performance for multi-class diabetic retinopathy detection. The curves confirm the model's capability to generalize well across varying disease severities.

Table 7: Comparison with Existing DR Detection Models

| Authors | Models | Accuracy |
|---|---|---|
| Amin Valizadeh et al. (2021) [22] | CNN | 83.30% |
| Chen, PN. Et al. (2021) [23] | NASNet-Large deep CNN | 87.50% |
| Shi B. et al. (2022) [24] | GoogleNet | 87.23% |
| R. Y. Sarobin et al. (2022) [25] | CNN | 75.61% |
| Kurup, G. et al. (2021) [26] | InceptionV-3 | 82.00% |
| **Proposed** | **ResNet18** | **81.37** |
| **Proposed** | **ResNet50** | **85.87** |
| **Proposed** | **ResNet152** | **90.15** |

Table 7 shows the comparative analysis of the proposed and existing DR detection and classification models based on their accuracy values. Amin Valizadeh et al. (2021) achieved an accuracy of 83.30% with CNN, while Chen et al. (2021) and Shi et al. (2022) employed NASNet-Large deep CNN and GoogleNet, respectively, achieving higher accuracies of 87.50% and 87.23%. R. Y. Sarobin et al. (2022) reported a lower accuracy of 75.61% using CNN, and Kurup et al. (2021) achieved 82.00% with InceptionV-3. In contrast, the proposed models—ResNet18, ResNet50, and ResNet152—demonstrated progressively higher accuracies of 81.37%, 85.87%, and 90.15%, respectively, showcasing the improved performance of these advanced architectures over the existing models.

**Conclusion and Future scope**

Diabetic Retinopathy is a progressive retinal disorder caused by prolonged diabetes and remains one of the leading causes of preventable blindness worldwide. Timely and accurate classification of DR stages is critical to initiating early treatment and mitigating vision loss. However, manual screening methods are often labor-intensive, subjective, and inaccessible in rural or resource-limited regions.

This study proposed a deep learning-based framework that integrates Canny Edge Detection for retinal structure segmentation, MobileNetV2 for lightweight feature extraction, and ResNet variants (ResNet-18, ResNet-50, ResNet-152) for classifying DR into five severity levels. The use of Canny Edge segmentation helped enhance vessel and lesion boundaries without relying on large manually annotated datasets, while MobileNetV2 effectively extracted discriminative features with minimal computational overhead. Among the three ResNet architectures evaluated, ResNet-152 demonstrated superior performance with a validation accuracy of 90.15%, F1-score of 90.01%, and an ROC-AUC of 0.94, significantly outperforming traditional CNN models and other baseline approaches. Class-wise evaluation also revealed high sensitivity and specificity across all DR stages, especially for challenging categories like Moderate and Proliferative DR.

The proposed Dl model can be extended to integrate real-time DR monitoring systems for clinical environments. Incorporating multi-modal data, such as patient history and genetic predispositions, could improve classification accuracy further. Exploring advanced deep learning techniques like attention mechanisms and transformer-based architectures may enhance the model's capability to identify subtle retinal abnormalities. Additionally, deploying the model on edge devices for remote screening in underserved regions could significantly improve healthcare accessibility.

**References**

[1] Trucillo, P., & Di Maio, E. (2021). Classification and Production of Polymeric Foams among the Systems for Wound Treatment. Polymers, 13(10), 1608. https://doi.org/10.3390/polym13101608

[2] Kropp M, Golubnitschaja O, Mazurakova A, Koklesova L, Sargheini N, Vo TKS, de Clerck E, Polivka J Jr, Potuznik P, Polivka J, Stetkarova I, Kubatka P, Thumann G. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications-risks and mitigation. EPMA J. 2023 Feb 13;14(1):21-42. doi: 10.1007/s13167-023-00314-8. PMID: 36866156; PMCID: PMC9971534.

[3] Tinajero MG, Malik VS. An update on the epidemiology of type 2 diabetes: a global perspective. Endocrinol Metab Clin North Am. 2021;50:337–355. doi: 10.1016/j.ecl.2021.05.013.

[4] Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis

from 1990. Sci Rep. 2020;10:14790. doi: 10.1038/s41598-020-71908-9.

[5] Hussain N, Edraki M, Tahhan R, Sanalkumar N, Kenz S, Akasha NK, Mtemererwa B, Mohammed N. Telemedicine for diabetic retinopathy screening using an ultra-widefield fundus camera. Clin Ophthalmol. 2017 Aug 14;11:1477-1482. doi: 10.2147/OPTH.S135287. PMID: 28860696; PMCID: PMC5565372.

[6] Shafi S, Parwani AV. Artificial intelligence in diagnostic pathology. DiagnPathol. 2023 Oct 3;18(1):109. doi: 10.1186/s13000-023-01375-z. PMID: 37784122; PMCID: PMC10546747.

[7] Kaur K, Gurnani B, Nayak S, Deori N, Kaur S, Jethani J, Singh D, Agarkar S, Hussaindeen JR, Sukhija J, Mishra D. Digital Eye Strain- A Comprehensive Review. Ophthalmol Ther. 2022 Oct;11(5):1655-1680. doi: 10.1007/s40123-022-00540-9. Epub 2022 Jul 9. PMID: 35809192; PMCID: PMC9434525.

[8] S. Kulkarni, N. Seneviratne, M. S. Baig, and A. H. A. Khan, "Artificial intelligence in medicine: where are we now?" Academic Radiology, vol. 27, no. 1, pp. 62–70, 2020.

[9] Y. Mori, H. Neumann, M. Misawa, S. E. Kudo, and M. Bretthauer, "Artificial intelligence in colonoscopy: now on the market. What's next?" Journal of Gastroenterology and Hepatology, vol. 36, 2021.

[10] H. Endo, S. Kase, M. Saito et al., "Choroidal thickness in diabetic patients without diabetic retinopathy: a metametaanalysis," American Journal of Ophthalmology, vol. 218, pp. 68–77, 2020.

[11] S. Roy and D. Kim, "Retinal capillary basement membrane thickening: role in the pathogenesis of diabetic retinopathy," Progress in Retinal and Eye Research, vol. 82, Article ID 100903, 2020.

[12] L. Camara Neto, G. L. B. Ramalho, J. F. S. Rocha Neto, R. M. S. Veras, and F. N. S. Medeiros, "An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images," Expert Systems with Applications, vol. 78, pp. 182–192, 2017.

[13] R. Sundaram, R. KS, P. Jayaraman, and V. B, "Extraction of blood vessels in fundus images of retina through hybrid segmentation approach," Mathematics, vol. 7, no. 2, p. 169, 2019.

[14] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," IEEE Transactions on Medical Imaging, vol. 34, no. 9, pp. 1797–1807, 2015.

[15] L. C. Rodrigues and M. Marengoni, "Segmentation of optic disc and blood vessels in retinal images using wavelets, mathematical morphology and Hessian-based multi-scale filtering," Biomedical Signal Processing and Control, vol. 36, pp. 39–49, 2017.

[16] Ç. Sazak, C. J. Nelson, and B. Obara, "+e multiscale bowlerhat transform for blood vessel enhancement in retinal images," Pattern Recognition, vol. 88, pp. 739–750, 2019.

[17] B. Savelli, A. Bria, A. Galdran et al., "Illumination correction by dehazing for retinal vessel segmentation," in Proceedings of the 30th International Symposium on Computer-Based Medical Systems, pp. 219–224, +essaloniki, Greece, June 2017.

[18] F. Girard, C. Kavalec, and F. Cheriet, "Joint segmentation and classification of retinal-arteries/veins from fundus images," Artificial Intelligence in Medicine, vol. 94, pp. 96–109, 2019.

[19] K. Hu, Z. Zhang, X. Niu et al., "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," Neurocomputing, vol. 309, pp. 179–191, 2018.

[20] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong, and J. Liu, "DeepVessel: retinal vessel segmentation via deep learning and conditional random field," in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 132–139, Athens, Greece, October 2016.

[21] T. A. Soomro, A. J. Afifi, J. Gao et al., "Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network," in Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, pp. 1–8, Sydney, Australia, November 2017.

[22] Valizadeh, Amin, Jafarzadeh Ghoushchi, Saeid, Ranjbarzadeh, Ramin, Pourasad, Yaghoub, Presentation of a Segmentation Method for a Diabetic Retinopathy Patient's Fundus Region Detection Using a Convolutional Neural Network,

Computational Intelligence and Neuroscience, 2021, 7714351, 14 pages, 2021. https://doi.org/10.1155/2021/7714351

[23] Chen, PN., Lee, CC., Liang, CM. et al. General deep learning model for detecting diabetic retinopathy. BMC Bioinformatics 22 (Suppl 5), 84 (2021). https://doi.org/10.1186/s12859-021-04005-x

[24] Shi, B.; Zhang, X.;Wang, Z.; Song, J.; Han, J.; Zhang, Z.; Toe, T.T. GoogLeNet-based Diabetic-Retinopathy-Detection. In Proceedings of the 14th IEEE International Conference on Advanced Computational Intelligence (ICACI),Wuhan, China, 15–17 July 2022; pp. 246–249.

[25] R., Y.; Sarobin, M.V.R.; Panjanathan, R.; Jasmine, S.G.; Anbarasi, L.J. Diabetic Retinopathy Classification Using CNN and Hybrid Deep Convolutional Neural Networks. Symmetry 2022, 14, 1932

[26] Kurup, G.; Jothi, J.A.A.; Kanadath, A. Diabetic Retinopathy Detection and Classification using Pretrained Inception-v3. In Proceedings of the IEEE International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Pune, India, 29–30 October 2021; pp. 1–6.