Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 25, No. 1, November 2025, pp. 161∼172
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 10/C/C3/DT.05.00/2025
DOI: 10.30812/matrik.v25i1.4976

Proliferative Diabetic Retinopathy Detection Using Convolutional Neural Network with Enhanced Retinal Image

Wilda Imama Sabilla, Ariadi Retno Tri Hayati Ririd, Mamluatul Hani'ah, Astrifidha Rahma Amalia
Politeknik Negeri Malang, Malang, Indonesia

Article Info

Article history:

Received March 11, 2025
Revised September 09, 2025
Accepted November 07, 2025

Keywords:

Convolutional Neural Network;
Diabetic Retinopathy;
Image Enhancement;
PDR Detection;
Retinal Image.

ABSTRACT

Proliferative Diabetic Retinopathy (PDR) is the most severe stage of Diabetic Retinopathy (DR), carrying the highest risk of complications. Automatic detection can help provide earlier and more accurate PDR diagnosis, but prediction accuracy may decline due to limitations in retinal images. Therefore, image enhancement techniques are often applied to improve DR classification. This study aims to detect PDR from retinal images using Convolutional Neural Networks (CNNs) and to evaluate the impact of three enhancement methods. This research method is based on a CNN architecture, including ResNet34, InceptionV2, and DenseNet121, as well as enhancement methods such as CLAHE, Homomorphic Filtering (HF), and Multiscale Contrast Enhancement (MCE). The results of this research show that CNN performance varies across architectures and enhancement methods. The highest performance was achieved using ResNet34 with HF, yielding an accuracy of 0.976, precision of 0.934, and recall of 0.904. CLAHE generally improved performance across architectures, achieving the best average accuracy of 0.953, whereas MCE decreased classification accuracy. Overall, the findings highlight the importance of selecting appropriate enhancement methods to improve PDR detection accuracy. Implementing such systems in clinical screening could help reduce the risk of vision impairment among diabetic patients.

Corresponding Author:

Wilda Imama Sabilla, +6285649655285,
Department of Information Technology,
Politeknik Negeri Malang, Malang, Indonesia,
Email: wildaimama@polinema.ac.id

How to Cite:
Wilda Imama Sabilla, "Proliferative Diabetic Retinopathy Detection Using Convolutional Neural Network with Enhanced Retinal Image", *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 25, No. 1, pp. 161-172, November, 2025.

1. INTRODUCTION

Diabetes is a condition that disrupts insulin production, a hormone that regulates blood sugar levels and converts glucose into energy. Diabetes also results in disturbances in the metabolism of carbohydrates, fats, and proteins [1]. According to the 2021 report from the Ministry of Health of the Republic of Indonesia, 19.47 million people were living with diabetes in Indonesia [2]. The 2022 report from the International Diabetes Federation (IDF) stated that Indonesia ranked first in ASEAN for the number of type 1 diabetes patients, with a total of 41,817 cases [3]. Diabetic retinopathy, also known as diabetic eye disease, is a condition that damages the retina in individuals with diabetes. The number of diabetic retinopathy patients reached 126.6 million in 2010 and is projected to rise to 191 million by 2030 [4]. To detect this disease, a doctor must manually identify it, which is a lengthy process. Therefore, a method is needed to assist and facilitate ophthalmologists in diagnosing diabetic retinopathy more quickly and accurately in retinal images [5]. Diabetic retinopathy (DR) is divided into five stages: mild, moderate, severe, proliferative, and normal (without diabetic retinopathy) [6]. Proliferative Diabetic Retinopathy (PDR) represents the most advanced and vision-threatening stage of diabetic retinopathy. Structural changes in retinal blood vessels occur in the PDR stage. Fragile and easily ruptured new blood vessels form on the retina of PDR patients, which can cause permanent vision loss if not treated immediately. Compared to earlier stages of DR, PDR has a much higher risk of severe complications and requires more cost for treatment [7]. Thus, early detection of PDR is essential to prevent irreversible vision loss, minimize treatment costs, and improve patients' quality of life.

Machine learning is widely used for research in various fields. Deep learning is a subfield of machine learning that utilizes deep neural network architectures to generate predictions or classifications. One type of deep learning is the Convolutional Neural Network. Convolutional Neural Network is often used in image classification, object detection, and object tracking [8]. Convolutional Neural Networks have been applied in various fields, including skin cancer pigment classification [9], wood types [10], crowd detection [11], and food classification [12]. CNNs are widely used for medical image classification because they can handle complex data and automatically extract features from medical images. The application of CNN in medical imaging includes the detection of DR in retinal images. Image enhancement methods are often applied to improve the quality of the images to be processed. Several studies have demonstrated the effectiveness of CNNs for diabetic retinopathy detection, and they are often combined with image enhancement techniques to improve classification accuracy. Qiao et al. [13] focused on early-stage detection, which is beneficial for prevention but less effective in addressing advanced stages such as Proliferative Diabetic Retinopathy (PDR). Similarly, Putra et al. [14] examined Non-Proliferative DR (NPDR) detection with image enhancement techniques such as CLAHE, Morphology Contrast Enhancement, and Homomorphic filtering combined with CNN architectures like GoogLeNet and ResNet variants. While Hemanth et al. [15] applied CNN combined with Histogram Equalization (HIST) and Contrast Limited Adaptive Histogram Equalization (CLAHE) for DR diagnosis, and Hayati et al. [5] investigated DR detection using CLAHE and various CNN architectures, including VGG16, InceptionV3, and EfficientNetB4.

Further, Alwakid et al. incorporated CNN DenseNet-121 [16] and Inception-V3 [17] architectures with enhancement techniques, including Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN), HIST, and CLAHE, to differentiate the stage of DR. El-Hoseny [18] utilized VGG16 for distinguishing between different classes of diabetes, while Ishtiaq et al. [19] combined machine learning–based feature extraction with CNN for DR stage classification. Pamungkas et al. [20] explored multiple architectures such as EfficientNet-B4, ResNet-50, DenseNet-201, Xception, and Inception-ResNet-v2, and Lee et al. [21] modified VGG-16 and ResNet-50 with dropout for severity grading, using CLAHE during preprocessing to improve image quality. Previous work demonstrated that CNNs with various architectures can achieve high accuracy in classifying DR stages, particularly early-stage (non-proliferative) stages. Their research also showed that applying image enhancement methods, such as CLAHE and other preprocessing techniques, can improve classification performance. There are gaps that previous research has not resolved: most existing studies did not systematically compare different enhancement methods nor specifically evaluate their effectiveness for PDR, which is the most vision-threatening stage due to the formation of fragile new blood vessels and the high risk of permanent vision loss. This creates a critical research gap, as accurate detection of PDR is essential for timely intervention and prevention of blindness.

The novelty of this research lies in its specific focus on detecting PDR using CNNs combined with retinal image enhancement techniques. The purpose of this study is to develop and evaluate a CNN-based model for detecting PDR in enhanced retinal images, thereby improving classification accuracy and providing a reliable tool for clinical decision-making. This problem is important to address because PDR involves fragile new blood vessels that can rupture and cause irreversible vision loss if left untreated promptly. Accurate detection of PDR enables timely intervention, reduces healthcare costs, and helps prevent blindness among diabetic patients. In this study, we evaluate three image enhancement techniques-CLAHE, homomorphic filtering, and multiscale contrast enhancement-with CNN architectures, including ResNet34, InceptionV2, and DenseNet121, to identify the most effective techniques for PDR detection.

2. RESEARCH METHOD

In general, the proposed method is illustrated in Figure 1, with retinal images serving as the input data. The preprocessing stage is carried out sequentially: first, the green channel is extracted, and then the images are resized to ensure uniformity. Following preprocessing, image enhancement is performed separately using three different techniques: CLAHE, Homomorphic Filtering, and Multiscale Contrast Enhancement. Each enhanced image is then independently subjected to feature extraction and classification using a Convolutional Neural Network (CNN). For this study, the APTOS 2019 dataset is used, which categorizes images into two classes: normal and PDR. The ratio of normal to PDR classes in the dataset is approximately 6:1.

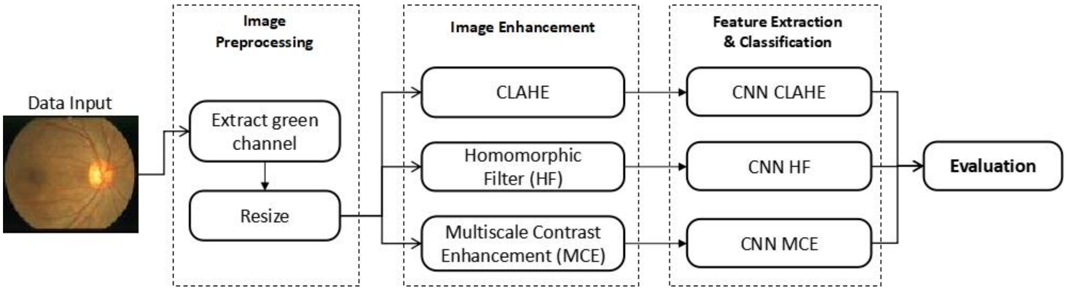

Figure 1. Research method flow

2.1. Data Input

The input data consists of retinal images in RGB format, sourced from the Kaggle APTOS 2019 dataset [22]. The APTOS 2019 dataset was selected because it provides a larger number and greater variety of retinal images compared to MESSIDOR 1 and 2, offers better image quality and more consistent class labeling than EyePACS, and is publicly accessible. This dataset contains 3,662 retinal images. Image resolution varies from 1,536 $\times$ 2,048 to 2,848 $\times$ 4,288, depending on the device used to capture the images. While the images are high resolution, some may appear blurry, overexposed, or underexposed. Additionally, certain images contain noise, blurring, and lighting variations, which can impact the model's accuracy. To address these issues, image enhancement techniques, including CLAHE, homomorphic filtering, and multiscale contrast enhancement, were applied to improve image quality before classification.

Classes with higher severity levels tend to have fewer samples compared to the "no retinopathy" class. An example image from the APTOS 2019 dataset is shown in Figure 2. In this dataset, the classes are divided into five categories: 0 - No DR, 1 - Mild, 2 - Moderate, 3-Severe, and 4 - Proliferative DR. Since this study focuses on detecting proliferative diabetic retinopathy (PDR), data selection is performed before analysis. Only data labeled as 0 and 4 are included in the study, while data with other labels are removed. A total of 2,100 data points are used as input, comprising 1,805 in class 0 (normal) and 295 in the PDR class. In this study, no specific balancing techniques were applied. The model was trained directly on the imbalanced data distribution to evaluate the effects of image enhancement and CNN architectures on PDR detection. This approach also reflects the actual prevalence of PDR in clinical settings.

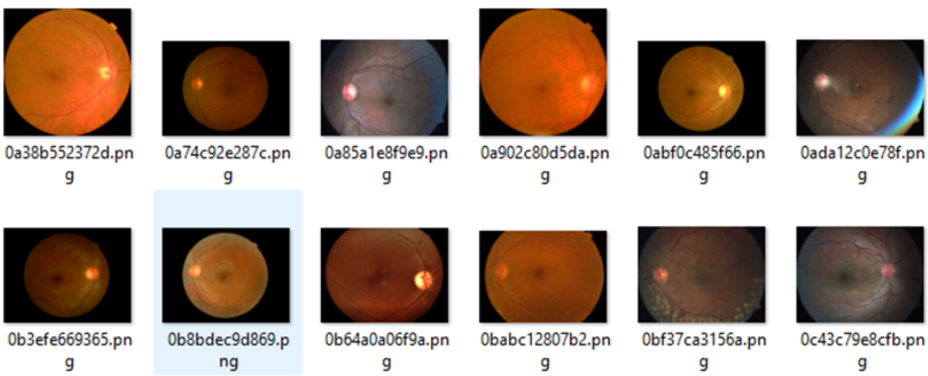

Figure 2. Example of an image in the APTOS 2019

2.2. Preprocessing

During preprocessing, two sequential steps were applied: extraction of the green channel to emphasize retinal structures, followed by image resizing to a standardized dimension. Using the green channel in retinal images is considered more effective because it offers higher contrast than the red or blue channels, making features such as microaneurysms, hemorrhages, and exudates more visible. This enhanced visibility is crucial for identifying diabetic retinopathy, which is often challenging to detect in other color channels [23]. The original image and the green-channel image are shown in Figure 3.

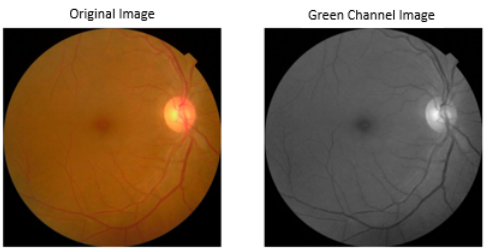

Figure 3. Original image and green channel image

The next data preprocessing step involves resizing the images to ensure uniform dimensions. Image resizing is the process of adjusting an image's size to match the input dimensions required by a Convolutional Neural Network (CNN) model. High-resolution images demand significant computational resources, both in terms of memory and processing time. By resizing the images, we can reduce their dimensions while preserving the essential information needed for classification. After resizing, the image dimensions used as input for the CNN are (224×224) pixels. The input size of 224×224 pixels was chosen because many CNN architectures are pretrained on ImageNet with this input resolution [24]. Additionally, previous work in medical imaging (such as breast ultrasound classification) has shown that 224×224 often strikes a good balance in preserving detail for feature extraction while keeping computational costs manageable [25]. The image before resizing and the image after resizing are shown in Figure 4.

Figure 4. Before and after resizing

2.3. Image Enhancement

Image enhancement methods are techniques used to improve image quality or extract valuable information from it. The primary goal is to make the image easier to interpret, enhance visual features, or reduce noise. In this study, image enhancement techniques-CLAHE, homomorphic filtering, and multiscale contrast enhancement-were applied separately to retinal images prior to PDR detection using a CNN. We selected CLAHE, homomorphic filtering, and multiscale contrast enhancement as image enhancement techniques because each method offers complementary benefits for medical images. CLAHE is well-suited for enhancing local contrast while limiting noise amplification, as has been demonstrated by Hayati et al. [5]. Homomorphic filtering is effective in correcting nonuniform illumination and enhancing high-frequency structural detail [14]. Meanwhile, multiscale contrast enhancement amplifies structures across multiple spatial scales and is closely related to superpixel-based multiscale homomorphic techniques that preserve both fine and coarse features. By comparing these three methods, we aim to determine which image enhancement approach best supports CNN performance in detecting Proliferative Diabetic Retinopathy.

1. CLAHE
CLAHE (Contrast Limited Adaptive Histogram Equalization) is an image processing technique designed to enhance image contrast while preserving fine details and preventing excessive contrast in areas with high intensity variation. The CLAHE technique enhances contrast by processing small sections of an image, known as tiles. Each tile's contrast is adjusted to align its histogram with a predefined distribution. To ensure smooth transitions between tiles, bilinear interpolation is applied to blend adjacent regions [5]. This method aims to create a seamless overall appearance. Generally, the process includes calculating local histograms, determining cumulative distributions, and normalizing the image.

In this study, CLAHE was applied after resizing the image to 224 $\times$ 224 pixels. Several parameters are configured when applying CLAHE using the Python library. The parameter that controls the degree of contrast limiting in each local histogram block in CLAHE (ClipLimit) is set to 4.0. And the parameter that determines the grid size for local histogram equalization in CLAHE (tileGridSize) is set to (8, 8). The ClipLimit and the tileGridSize values were selected based on preliminary experiments, as they provided a balance between enhancing vessel visibility and avoiding over-amplification of noise. The image before and after applying CLAHE are shown in Figure 5.

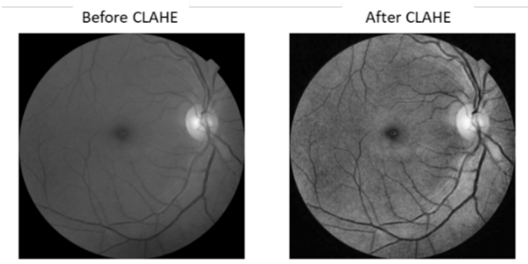

Figure 5. Before and after CLAHE

2. Homomorphic Filter
Homomorphic filtering is an image processing technique that enhances contrast while preserving both low- and high-intensity information. The method involves several steps. First, the input image is logarithmically transformed to reduce intensity variations. Next, the image is converted into the frequency domain using the Fourier Transform. In the frequency domain, high-frequency and low-frequency components are processed separately. Low-frequency components are adjusted to enhance contrast, while high-frequency components are modified to improve clarity. The processed image is then returned to the spatial domain through an inverse Fourier Transform [26].
Three parameters are configured for the homomorphic filter in the application using the Python library. The parameter d0 (cut-off frequency) determines the boundary between low-frequency components (illumination) and high-frequency components (details), with d0 = 40. The parameter gammaL (low-frequency gain) controls the attenuation of low-frequency components (illumination). The parameter gammaH (high-frequency gain) is a gain factor that controls the amplification of high-frequency components (details or reflectance). In this study, gammaL = 2.0 and gammaH = 2.5 are set. Homomorphic filtering was applied in the frequency domain using a Gaussian high-pass filter with a cutoff frequency of 40 and gain constants (γL = 2.0, γH = 2.5), based on some experiments, to reduce non-uniform illumination and enhance fine details. The images before and after applying homomorphic filtering are shown in Figure 6.

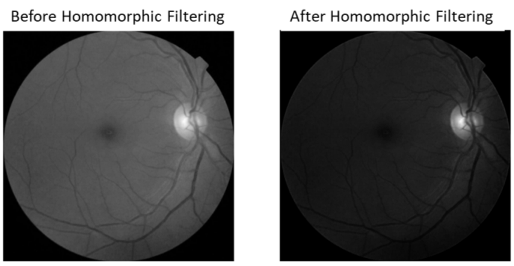

Figure 6. Before and after Homomorphic Filtering

3. Multiscale Contrast Enhancement
Multiscale Contrast Enhancement is an image processing technique that improves contrast across multiple scales or resolution levels. This method utilizes the wavelet transform, which efficiently captures features at both large and small scales. At each level of transformation, the image is divided into components—horizontal, vertical, and diagonal details—each of which can have its contrast enhanced. The wavelet transform is also recognized as one of the most effective methods for reducing random noise in images [27]. Two parameters set during multiscale contrast enhancement are the wavelet and the level. A wavelet is a mathematical function used to analyze signals or images in the time-frequency domain. In this study, the Python library was used, and the 'db1' (Daubechies level 1) wavelet was employed. The level parameter determines the number of scale decompositions applied to the image, with a value of 2. The images before and after applying multiscale contrast enhancement are shown in Figure 7.

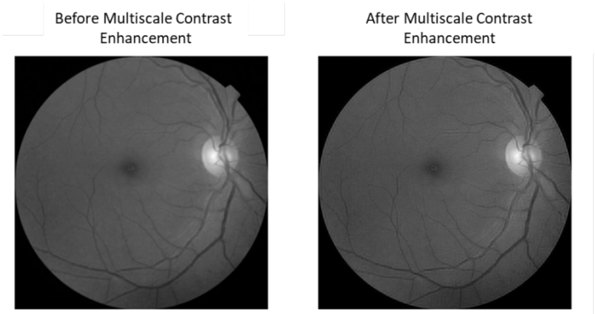

Figure 7. Before and after multiscale contrast enhancement

2.4. Feature Extraction and Classification

This stage extracts image features and classifies them as either normal or PDR using a CNN. A Convolutional Neural Network (CNN) is a specialized neural network architecture designed to process and interpret image data. It excels at identifying spatial structures and hierarchical features within visual inputs. In CNN, feature extraction is performed through the following stages:

a. Convolutional Layers: CNN uses convolutional layers to extract features from images. These layers contain filters, or kernels, that move across the image to detect specific patterns, such as edges, corners, or textures.
b. Pooling Layers: After passing through the convolutional layer, a pooling layer is applied to reduce the spatial dimensions of the extracted features. Max pooling or average pooling is used to retain the most important information by selecting the maximum or average value within a given region.
c. Activation Functions: The activation function, using ReLU (Rectified Linear Unit), is applied to introduce non-linearity into the model, allowing CNN to capture more complex relationships between extracted features.
d. Normalization Layers: CNN includes normalization layers (Batch Normalization) to accelerate convergence and ensure training stability.

Meanwhile, for classification in CNN, the stages used are as follows:

1. Flattening: After the convolutional and pooling layers, the output is flattened into a one-dimensional vector. This process organizes the extracted features into a format that fully connected layers can process.
2. Fully Connected Layers: These layers are connected to every neuron in the previous layer, allowing the network to understand relationships between extracted features and perform classification based on that information.
3. Activation Functions: Activation functions are applied in the output layer to generate a probability distribution for each class, enabling the model to predict the most likely class.
4. Loss Function: CNN uses a loss function to measure how close the model's predictions are to the actual class labels.
5. Optimization: The optimization process is used to adjust the weights and biases in the network to minimize the loss function. [28].

Three CNN architectures are used in this study: ResNet34, InceptionV2, and DenseNet121. A consistent set of parameter values is applied across all architectures during model training. The training parameters were selected based on both common practices in CNN-based medical image classification and empirical testing. Several experiments were conducted with varying batch sizes, epochs, and learning rates. The final configuration consisted of a batch size of 16, 10 epochs, and the Adam optimizer with a learning rate of 0.001. Binary cross-entropy was used as the loss function since the task involves binary classification. This configuration consistently produced the best balance between accuracy and training stability. This final configuration achieved the highest accuracy among the tested settings.

1. ResNet-34
   ResNet-34 is a member of the ResNet architecture family, introduced by He et al. in their paper "Deep Residual Learning for Image Recognition" [29]. This model incorporates residual blocks, enabling the network to grow deeper without suffering from the vanishing gradient issue commonly found in very deep networks. These residual blocks feature skip connections that bypass certain layers, allowing information to pass directly from input to output without traversing every intermediate layer. ResNet-34 is composed of 34 layers, utilizing 3×3 convolutions, pooling operations, and skip connections. Hayati et al [5] also used ResNet-34 for diabetic retinopathy classification.
2. InceptionV2
   InceptionV2 is an improved version of the Inception architecture, introduced by Szegedy et al. in 2015 [30]. It builds on InceptionV1 (GoogLeNet) and adds enhancements to improve computational efficiency. The Inception block enables the network to apply convolutional filters of various sizes and perform pooling simultaneously. To optimize performance, larger convolutions such as 5×5 are replaced with two consecutive 3×3 convolutions, reducing the number of parameters and accelerating computation. Additionally, batch normalization is integrated throughout the architecture to enhance training stability. Studies on MRI data for brain tumors show that InceptionV2 can be effective in medical imaging [31].
3. DenseNet121
   DenseNet was proposed by Gao Huang et al. in 2016 [32]. It features a unique connectivity pattern where each layer receives input from all preceding layers. DenseNet-121, a specific variant with 121 layers, utilizes dense blocks in which the output of each layer is directly linked to every subsequent layer within the same block. This design facilitates efficient information and gradient flow across the network. The architecture of DenseNet-121 comprises four dense blocks and incorporates batch normalization, ReLU activation, and convolutional layers throughout. DenseNet121 with a specific optimizer achieves high accuracy, demonstrating that this architecture is effective and competitive in DR classification [33].

2.5. Model Evaluation

To evaluate the model, the metrics used are accuracy, precision, and recall. Normal data predicted as normal is counted as a True Positive (TN), whereas data predicted as PDR is considered a False Negative (FN). Similarly, PDR data predicted as PDR is counted as True Negative (TP), but if it is predicted as normal, it is considered False Positive (FN) [34]. Thus, accuracy, precision, and recall can be calculated using equations (1) to (3).

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$precision = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

Table 1. Test Scenario

Scenario	Enhancement Image Method	CNN Architecture
1	No	ResNet34
2	No	InceptionV2
3	No	DenseNet121
4	CLAHE	ResNet34

(continued on next page)

Table 2 (continued)

Scenario	Enhancement Image Method	CNN Architecture
5	CLAHE	InceptionV2
6	CLAHE	DenseNet121
7	Homomorphic filter	ResNet34
8	Homomorphic filter	InceptionV2
9	Homomorphic filter	DenseNet121
10	Multiscale contrast enhancement	ResNet34
11	Multiscale contrast enhancement	InceptionV2
12	Multiscale contrast enhancement	DenseNet121

Table 1 presents the testing scenarios, comparing classification evaluation results for PDR detection across various image enhancement techniques and different CNN architectures. The data used in the testing process comes from the APTOS 2019 dataset. The evaluation of the CNN model was conducted by splitting the input data into training and test datasets at a 80:20 ratio. The details of the data distribution are shown in Table 2.

Table 2. Classification Data Descriptions

Data Label	Training Data	Testing Data
Normal	1448	357
PDR	232	63
Total	1680	420

3. RESULT AND ANALYSIS

The CNN training process used ResNet34, InceptionV2, and DenseNet121 architectures. The dataset was split into training and test sets, and training was performed with a batch size of 16, a learning rate of 0.001, the Adam optimizer, and binary cross-entropy loss. The models were trained for 10 epochs. During training, the accuracy gradually increased while the loss decreased, indicating that the networks successfully learned feature representations of retinal images. The testing phase was then performed on 420 images, comprising 357 normal and 63 PDR cases. Testing was conducted in 12 scenarios to evaluate system performance using accuracy, precision, and recall metrics. PDR classification testing was performed by modifying the test images, including original retinal images and retinal images enhanced with CLAHE, homomorphic filtering, and multiscale contrast enhancement. Then CNN architectures were used, namely ResNet34, InceptionV2, and DenseNet121.

The evaluation results for the original images without enhancement are shown in Table 3. In this table, the classification accuracy is relatively high across all architectures. The highest accuracy was achieved with the ResNet34 architecture at 0.959. For precision, the highest value of 0.916 was obtained using the InceptionV2 architecture. However, InceptionV2 had the lowest recall, 0.698, meaning that out of 63 PDR images in the test set, only 44 were correctly detected, while 19 were misclassified as normal.

Table 3. The Evaluation Results for The Original Images

Architecture	Accuracy	Precision	Recall	TP (PDR detected)	FN (PDR missed)
ResNet34	0.959	0.896	0.825	52	11
InceptionV2	0.945	0.916	0.698	44	19
DenseNet121	0.935	0.764	0.825	52	11
Average	0.946	0.859	0.783	49.3	13.7

For images enhanced with CLAHE, the evaluation results are shown in Table 4. An improvement in classification accuracy is observed for the DenseNet121 and InceptionV2 architectures, whereas it remains unchanged for ResNet34. The highest accuracy was achieved with the ResNet34 architecture at 0.959. With CLAHE, InceptionV2 achieved a recall of 0.825, detecting 52 PDR cases correctly and missing 11. Additionally, the average precision increased.

Table 4. The Evaluation Results for CLAHE-Enhanced Images

Architecture	Accuracy	Precision	Recall	TP (PDR detected)	FN (PDR missed)
ResNet34	0.959	0.896	0.825	52	11
InceptionV2	0.957	0.881	0.825	52	11
DenseNet121	0.943	0.898	0.698	44	19
Average	0.953	0.892	0.783	49.3	13.7

Table 5 presents the evaluation results for images enhanced with the homomorphic filter. The table shows an improvement in classification accuracy for the ResNet34 and InceptionV2 architectures. However, accuracy, precision, and recall values decreased when using the DenseNet121 architecture. The highest accuracy was achieved with the ResNet34 architecture at 0.976, meaning that 57 of 63 PDR cases were correctly detected while only six were missed. There was also an increase in average precision, but the recall for DenseNet121 was significantly low.

Table 5. The Evaluation Results for The Homomorphic Filter-Enhanced Images

Architecture	Accuracy	Precision	Recall	TP (PDR detected)	FN (PDR missed)
ResNet34	0.976	0.934	0.904	57	6
InceptionV2	0.964	0.913	0.841	53	10
DenseNet121	0.885	0.800	0.317	20	43
Average	0.942	0.882	0.687	43.4	19.6

The evaluation results for images enhanced with multiscale contrast enhancement are presented in Table 6. An improvement in classification accuracy is observed for the ResNet34 and InceptionV2 architectures. However, accuracy, precision, and recall values decreased for the DenseNet121 architecture, although the decline was not as significant as with the homomorphic filter. The highest accuracy was achieved with the ResNet34 architecture at 0.962. With ResNet34, recall reached 0.841, meaning 53 PDR cases were correctly detected, while 10 were missed. Additionally, the average precision increased.

Table 6. The Evaluation Results for Multiscale Contrast Enhancement Images

Architecture	Accuracy	Precision	Recall	TP (PDR detected)	FN (PDR missed)
ResNet34	0.962	0.898	0.841	53	10
InceptionV2	0.959	0.883	0.841	53	10
DenseNet121	0.902	0.805	0.460	29	34
Average	0.941	0.862	0.714	45	18

Based on the testing results, the accuracy, precision, and recall values for the original images were already relatively high, although recall remained below 80%. The highest accuracy, precision, and recall were achieved with ResNet34 and the homomorphic filter, at 0.976, 0.934, and 0.904, respectively. With the homomorphic filter, InceptionV2 also achieved the highest evaluation scores compared to other image enhancement techniques. On average, CLAHE-enhanced images produced the best evaluation results, with average accuracy, precision, and recall of 0.953, 0.892, and 0.783, respectively. The lowest performance was observed with multiscale contrast enhancement. The number of PDR cases successfully detected (TP) and missed (FN) varies across architectures and enhancement methods. The best result was obtained with ResNet34 combined with the homomorphic filter, which detected 57 of 63 PDR cases (recall 0.904). In contrast, DenseNet121 consistently produced lower recall, particularly with the homomorphic filter (recall 0.317, only 20 detected cases).

This research finds that CNN architectures can classify PDR from retinal images, with performance varying across networks and image enhancement techniques. ResNet34 consistently outperformed other architectures, achieving the best balance of accuracy, precision, and recall. The homomorphic filter in particular provided the highest scores, reaching 0.976 accuracy, 0.934 precision, and 0.904 recall. From a clinical perspective, recall is a critical metric because it reflects the ability to identify PDR cases correctly. False negatives represent missed detections, which could delay diagnosis and treatment for patients at risk. As shown in Table 5, ResNet34 missed only 6 cases, whereas DenseNet121 failed to detect 43 of 63 PDR images, making it unsuitable for reliable PDR detection. The performance differences also highlight the role of image enhancement. Among the three enhancement techniques evaluated, CLAHE and homomorphic filtering consistently improved the classification results, particularly in terms of accuracy and precision. CLAHE generally improved model stability across architectures, yielding the best average performance. Homomorphic filtering provided the highest peak performance, particularly for ResNet34, though DenseNet121 suffered greatly under this condition.

The experimental results demonstrated that image enhancement significantly improves CNN performance in detecting Proliferative Diabetic Retinopathy (PDR).

Compared to previous studies on diabetic retinopathy detection, our results are consistent with findings that CNN-based methods can achieve high accuracy when combined with appropriate preprocessing. The study by Putra et al. [14] reported that applying image enhancement and additional feature reduction improved DR detection accuracy from 85% to 87.5%. Among the tested methods, homomorphic filtering achieved the best results. These findings are consistent with those of Hayati et al. [5], who observed performance improvements across almost all CNN architectures when CLAHE was applied. However, a slight difference was noted: in Hayati's work, applying CLAHE with ResNet34 did not improve accuracy, whereas in the present study, it did. This discrepancy may be attributed to differences in parameter settings when applying CLAHE and ResNet34, suggesting that careful parameter selection can significantly enhance CNN performance. And according to Lee's study [21], CLAHE likely improves local contrast in fundus images, which tends to help CNN architectures more reliably identify critical features, particularly for severity grading. This interpretation aligns with the present results, which show that several combinations of enhancement methods and CNN architectures reduced PDR detection errors (false negatives). On the other hand, multiscale contrast enhancement showed limited benefit and, in some cases, reduced recall, suggesting that its parameter settings may not have been optimal for PDR feature extraction. In some cases, performance even decreased compared to the original, unenhanced images, especially in recall. This may indicate that MCE introduces noise or alters key retinal structures, making it harder for CNN architectures to extract relevant features for PDR detection.

There are two implications of these results. First, they recommend CLAHE and homomorphic filtering as preprocessing steps for CNN-based PDR detection systems. These techniques not only improve classification performance but also help mitigate the negative impact of variations in image illumination and contrast, which are often found in retinal datasets. Second, the relatively low recall in certain scenarios emphasizes the importance of handling dataset imbalance to avoid missed PDR diagnoses, which are clinically critical. Future research could address this limitation by applying data augmentation or generative adversarial networks (GANs) to improve the sensitivity of CNN models to PDR cases. Data augmentation techniques, such as rotation and flipping, can be applied to input images to increase data diversity. To improve model performance, transfer learning from pre-trained models can also be used.

4. CONCLUSION

This study aimed to evaluate the effectiveness of image enhancement techniques, including CLAHE, homomorphic filtering, and multiscale contrast enhancement, for the the detection of Proliferative Diabetic Retinopathy (PDR) using several CNN architectures (ResNet34, InceptionV2, and DenseNet121). The evaluation was conducted using the APTOS 2019 dataset to obtain normal and PDR classes. The experimental results demonstrate that CLAHE and homomorphic filtering consistently improved CNN performance. The highest performance was achieved using ResNet34 with homomorphic filtering, yielding an accuracy of 0.976, precision of 0.934, and recall of 0.904. These findings indicate that appropriate image enhancement significantly improves model accuracy and recall, which are critical for reducing false-negative cases in PDR detection. In contrast, multiscale contrast enhancement did not improve performance and, in some cases, decreased recall, indicating that not all enhancement methods are equally effective for PDR detection. The findings also highlight that class imbalance remains a challenge, as recall values for PDR detection were lower than for normal classes.

Overall, this study highlights the importance of selecting appropriate image enhancement techniques in medical image classification. By systematically comparing enhancement methods, the results deepen understanding of how image quality affects CNN performance in PDR detection and lay a foundation for developing more robust, clinically applicable automated diagnostic tools. Future work may address class imbalance problems through techniques such as data augmentation or oversampling. Additionally, further adjustments to image enhancement parameters can be made to find optimal values that maximize the quality of retinal images.

5. ACKNOWLEDGEMENTS

The authors sincerely appreciate Politeknik Negeri Malang for their invaluable support throughout this research project. Gratitude is also extended to the P2M reviewer team at Politeknik Negeri Malang for approving the research proposal. Furthermore, the researcher would like to thank the anonymous reviewers and editors for their insightful and constructive feedback, which has significantly improved the quality of this study.

6. DECLARATIONS

AUTHOR CONTIBUTION

The first author was responsible for designing the research flow, conducting the study, performing evaluations, and writing the research findings. The second author assisted in analysis and evaluation. The third author contributed to the evaluation and to the writing of the research findings. The fourth author assisted with data collection and the writing of the research findings.

FUNDING STATEMENT

Politeknik Negeri Malang supported this work under DIPA Fund Number: SP DIPA 023.18.2.677606/2024.

COMPETING INTEREST

The authors state that there are no competing interests.

REFERENCES

[1] L. Norton, C. Shannon, A. Gastaldelli, and R. A. DeFronzo, "Insulin: The master regulator of glucose metabolism," vol. 129, p. 155142, April,2022, https://doi.org/10.1016/j.metabol.2022.155142.

[2] J. Lemantara and T. Lusiani, "Analisis prediksi penyakit diabetes pada wanita menggunakan metode naïve bayes dan k-nearest neighbor," vol. 12, no. 3, August,2024, https://doi.org/10.23960/jitet.v12i3.4911.

[3] C. M. Annur. (2024) Ada 41 ribu penderita diabetes tipe 1 di indonesia pada 2022, terbanyak di ASEAN.

[4] I. Kandel and M. Castelli, "Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review," vol. 10, no. 6, p. 2021, March,2020, https://doi.org/10.3390/app10062021.

[5] M. Hayati *et al.*, "Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning," vol. 216, pp. 57–66, 2023, https://doi.org/10.1016/j.procs.2022.12.111.

[6] R. Schwartz *et al.*, "Objective evaluation of proliferative diabetic retinopathy using OCT," vol. 4, no. 2, pp. 164–174, February,2020, https://doi.org/10.1016/j.oret.2019.09.004.

[7] S. Vaz-Pereira, T. Morais-Sarmento, and R. Esteves Marques, "Optical coherence tomography features of neovascularization in proliferative diabetic retinopathy: A systematic review," vol. 6, no. 1, p. 26, December,2020, https://doi.org/10.1186/s40942-020-00230-3.

[8] K. Vijiyakumar, V. Govindasamy, and V. Akila, "An effective object detection and tracking using automated image annotation with inception based faster R-CNN model," vol. 5, pp. 343–356, 2024, https://doi.org/10.1016/j.ijcce.2024.07.006.

[9] L. Hakim, Z. Sari, and H. Handhajani, "Klasifikasi citra pigmen kanker kulit menggunakan convolutional neural network," vol. 5, no. 2, pp. 379–385, April,2021, https://doi.org/10.29207/resti.v5i2.3001.

[10] A. S. Oktaria, E. Prakasa, and E. Suhartono, "Wood species identification using convolutional neural network (CNN) architectures on macroscopic images," vol. 4, no. 3, pp. 274–283, December,2019, https://doi.org/10.25126/jitecs.201943155.

[11] M. Mentari, W. I. Sabilla, K. S. Batubulan, A. Latif, A. R. Fitriana, and Atmayanti, "Crowd counting during a pandemic to find out community response to activity restriction policy using deep learning," pp. 101–108, September,2022, https://doi.org/10.1109/IEIT56384.2022.9967905.

[12] A. Chaitanya, J. Shetty, and P. Chiplunkar, "Food image classification and data extraction using CNN and web crawlers," vol. 218, pp. 143–152, 2023, https://doi.org/10.1016/j.procs.2022.12.410.

[13] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," vol. 8, pp. 104 292–104 302, 2020, https://doi.org/10.1109/ACCESS.2020.2993937.

[14] R. Putra, H. Tjandrasa, and N. Suciati, "Severity classification of non-proliferative diabetic retinopathy using convolutional support vector machine," vol. 13, no. 4, pp. 156–170, August,2020, https://doi.org/10.22266/ijies2020.0831.14.

[15] D. J. Hemanth, O. Deperlioglu, and U. Kose, "Retracted Article: An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network," vol. 32, no. 3, pp. 707–721, February,2020, https://doi.org/10.1007/s00521-018-03974-0.

[16] G. Alwakid, W. Gouda, and M. Humayun, "Enhancement of diabetic retinopathy prognostication using deep learning, Clahe, and Esrgan," vol. 13, no. 14, p. 2375, July,2023, https://doi.org/10.3390/diagnostics13142375.

[17] ——, "Deep learning-based prediction of diabetic retinopathy using Clahe and Esrgan for enhancement," vol. 11, no. 6, p. 863, March,2023, https://doi.org/10.3390/healthcare11060863.

[18] H. M. El-Hoseny, H. F. Elsepae, W. A. Mohamed, and A. S. Selmy, "Optimized deep learning approach for efficient diabetic retinopathy classification combining VGG16-CNN," vol. 77, no. 2, pp. 1855–1872, 2023, https://doi.org/10.32604/cmc.2023.042107.

[19] U. Ishtiaq, E. R. M. F. Abdullah, and Z. Ishtiaque, "A hybrid technique for diabetic retinopathy detection based on ensemble-optimized CNN and texture features," vol. 13, no. 10, p. 1816, May,2023, https://doi.org/10.3390/diagnostics13101816.

[20] Y. Pamungkas, E. Triandini, W. Yunanto, and Y. Thwe, "Enhancing diabetic retinopathy classification in fundus images using CNN architectures and oversampling technique," vol. 6, no. 1, pp. 413–425, February,2025, https://doi.org/10.18196/jrc.v6i1.25331.

[21] C.-H. Lee and Y.-H. Ke, "Fundus images classification for diabetic retinopathy using deep learning," pp. 264–270, June,2021, https://doi.org/10.1145/3474963.3475849.

[22] M. Karthik and D. Sohier. (2025) Aptos 2019 blindness detection. [Online]. Available: https://kaggle.com/competitions/aptos2019-blindness-detection

[23] P. K. Jena, B. Khuntia, C. Palai, M. Nayak, T. K. Mishra, and S. N. Mohanty, "A novel approach for diabetic retinopathy screening using asymmetric deep learning features," vol. 7, no. 1, p. 25, January,2023, https://doi.org/10.3390/bdcc7010025.

[24] I. Al-Kamachy, R. Hassanpour, and R. Choupani, "Classification of diabetic retinopathy using pre-trained deep learning models," 2024.

[25] S. Tang *et al.*, "The effect of image resolution on convolutional neural networks in breast ultrasound," vol. 9, no. 8, p. e19253, August,2023, https://doi.org/10.1016/j.heliyon.2023.e19253.

[26] S. Adhikari and S. P. Panday, "Image enhancement using successive mean quantization transform and homomorphic filtering," pp. 1–5, November,2019, https://doi.org/10.1109/AITB48515.2019.8947437.

[27] K. Pyka, "Wavelet-based local contrast enhancement for satellite, aerial and close range images," vol. 9, no. 1, p. 25, January,2017, https://doi.org/10.3390/rs9010025.

[28] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," vol. 57, no. 4, p. 99, March,2024, https://doi.org/10.1007/s10462-024-10721-6.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, June,2016, pp. 770–778, https://doi.org/10.1109/CVPR.2016.90.

[30] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," 2015.

[31] A. A. S. Ali, "Brain tumor classification using a hybrid deep learning model: Leveraging densenet121 and inceptionv2 architectures," vol. 5, no. 4, pp. 455–463, December,2024, https://doi.org/10.47372/ejua-ba.2024.4.402.

[32] G. Huang, Z. Liu, p. d. u. family=Maaten, given=L., and K. Q. Weinberger, "Densely connected convolutional networks," 2018.

[33] T. Mustaqim, P. H. Safitri, and D. Muhajir, "A deep learning model comparation for diabetic retinopathy image classification," vol. 12, no. 1, pp. 21–30, April,2025, https://doi.org/10.15294/sji.v12i1.20939.

[34] W. I. Sabilla, C. B. Vista, and D. S. Hormansyah, "Implementasi multilayer perceptron untuk memprediksi harapan hidup pada pasien penyakit kardiovaskular," vol. 6, no. 1, pp. 57–68, March,2023.