JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

DIABETIC RETINOPATHY SEVERITY CLASSIFICATION USING GAMMA CORRECTION-BASED IMAGE ENHANCEMENT AND BN-VGG ARCHITECTURE

Indri Ramayanti[1*]; Karnadi[1]; Septiani Nadra Indawaty[1]; Muhammad Umar Abdussalam[1]; Malika Zilda[1]; Anita Desiani[2]

Faculty of Medicine, Faculty of Engineering[1]
Universitas Muhammadiyah Palembang, Palembang, Indonesia[1]
https://www.um-palembang.ac.id[1]
indri_ramayanti@um-palembang.ac.id*, karnadi@um-palembang.ac.id,
septianinadraiindawaty91@yahoo.com, umarabdussalam0@gmail.com, malikazilda09@gmail.com

Faculty of Mathematics and Natural Sciences[2]
Universitas Sriwijaya, Indralaya, Indonesia[2]
https://unsri.ac.id[2]
anita_desiani@unsri.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Diabetic retinopathy (DR) is a diabetes-related condition that can cause vision impairment or vision loss. Accurately identifying the level of DR from retinal fundus images is crucial for early detection. However, poor image quality often degrades classification performance. This study proposes an approach that integrates gamma correction-based image enhancement with a Batch Normalization–Visual Geometry Group (BN-VGG) architecture for multiclass DR severity classification. Gamma correction is applied to improve image contrast, while BN-VGG enhances training stability and feature representation. The proposed method categorizes DR into five classifications: normal, mild, moderate, severe, and proliferative. The enhanced images achieved PSNR of 30.85 and SSIM above 0.86, indicating improved visual quality. The model achieved accuracy at 0.97, sensitivity at 0.92, specificity at 0.98, F1-score at 0.92, Cohen's Kappa at 0.90, and G-Mean at 0.97. The innovative aspect of this study is the incorporation of gamma correction with BN-VGG architecture, demonstrating that image enhancement can significantly improve multiclass DR classification performance without increasing model complexity. The study's results indicate the proposed method's effectiveness for accurate & reliable DR severity classification.*

Keywords: *Batch Normalization, Classification, Diabetic Retinopathy, Gamma Correction, Visual Geometry Group.*

Intisari— *Diabetic Retinopathy (DR) merupakan komplikasi diabetes yang berpotensi menimbulkan gangguan penglihatan bahkan kebutaan. Identifikasi level keparahan DR secara tepat melalui citra fundus retina memiliki peran penting dalam proses deteksi dini. Namun, kualitas citra yang rendah sering menurunkan kinerja klasifikasi. Penelitian ini mengusulkan pendekatan yang mengintegrasikan peningkatan kualitas citra berbasis gamma correction dengan arsitektur Batch Normalization–Visual Geometry Group (BN-VGG) untuk klasifikasi multikelas tingkat keparahan DR. Gamma correction digunakan untuk meningkatkan kontras citra, sedangkan BN-VGG meningkatkan stabilitas pelatihan dan representasi fitur. Metode yang diusulkan mengklasifikasikan DR ke dalam lima kategori, yaitu normal, ringan, sedang, berat, dan proliferatif. Citra hasil peningkatan mencapai PSNR sebesar 30,85 dan SSIM di atas 0,86, yang menunjukkan peningkatan kualitas visual. Model menghasilkan akurasi sebesar 0,97, sensitivitas 0,92, spesifisitas 0,98, skor F1 0,92, Cohen's Kappa 0,90, dan G-Mean 0,97. Kebaruan penelitian ini terletak pada* *integrasi efektif antara gamma correction dan arsitektur BN-VGG, yang menunjukkan bahwa peningkatan kualitas citra dapat secara signifikan meningkatkan kinerja klasifikasi multikelas tanpa menambah kompleksitas model. Hasil ini menunjukkan bahwa metode yang diusulkan efektif untuk klasifikasi tingkat keparahan DR yang akurat dan andal.*

***Kata Kunci**: Batch Normalization, Klasifikasi, Retinopati Diabetik, Koreksi Gamma, Visual Geometry Group.*

INTRODUCTION

Diabetic retinopathy (DR) is diabetes-related condition that can cause vision impairment or vision loss in humans. This condition occurs when elevated glucose levels block blood vessels in the eye, causing swelling and blood leakage[1]. DR is difficult to detect in the early stages because it causes no symptoms in the patient, but in the advanced stages, the patient experiences visual disturbances such as blurred vision, spots, and loss of visual acuity [2]. Early identification is necessary to lower the risk of DR illness and avoid it in its later stages. An ophthalmologist manually examines retinal images, which are captured by a fundus camera, to diagnose DR disease early [3]. A manual examination is a time-consuming and costly process. It is also subjective, which can lead to misdiagnosis. The limitations of manual examination can be overcome with an automated system to detect DR disease, one of which uses image classification [4].

Image classification is a process used to identify the class or label of an object [5]. Image classification is performed by examining and extracting features from retinal images obtained from medical examinations. The performance of image classification is affected by the quality of the images [6]. Retinal fundus images are usually of low quality, with uneven illumination [7]. Low contrast in retinal images can be improved by applying contrast enhancement techniques. One of the quality enhancement techniques that can be used to enhance and equalize contrast is gamma correction. Gamma correction is a method of modifying image contrast and brightness by applying an exponential function to pixel intensity values [7], [8], [9]. This method uses the gamma value to adjust the brightness and contrast levels of the image to match human visual perception better and is able to preserve the details of the image features [10]. Xue et al. [11] applied gamma correction to diabetic retinopathy classification using Vmamba-m, resulting in 80% accuracy, but the F1-score produced was still below 60%. Aziz et al. [12] applied image quality improvement with gamma correction to diabetic retinopathy classification resulting in accuracy performance above 90%, but unfortunately the resulting precision is still below 80%. Image quality enhancement can improve image quality, resulting in better classification performance. Automatic image classification has been developed using deep learning methods.

Automatic learning of hierarchical representations of features is a key strength of deep learning models, which have demonstrated remarkable efficacy in medical image classification. One widely used architecture is the Visual Geometry Group (VGG) model, which consists of 16–19 layers. The architecture consists of convolutional, max pooling, and fully connected layers [13]. The VGG's depth enables effective feature extraction from images [14]. Abhishek and Jatin [15] and Nair et al. [16] applied VGG to diabetic retinopathy classification without image quality enhancement, resulting in relatively low accuracy below 78%. Furthermore, these studies did not comprehensively evaluate model performance using additional metrics. Houby [17] applied VGG to diabetic retinopathy classification resulted in an accuracy of 74% with other metrics like as recall and precision still below 70%. Even though deep learning architectures like VGG have shown great results in retinal image classification, many publicly available diabetic retinopathy datasets contain a limited number of labeled images. This limitation increases the potential for overfitting and may affect the model's generalizability. To address this challenge, data augmentation techniques are commonly applied to expand data variability and improve model robustness [5], [18], [19]. In this study, augmentation strategies including rotation, flipping, and median filtering are employed to mitigate data scarcity and support stable training. Nevertheless, generalization to broader clinical settings remains an important consideration.

VGG has a deep network architecture, which can cause the model to struggle with weight initialization. Weights that are not properly initialized can cause gradient problems, which can hinder weight updates and slow convergence during model training [20]. One technique that can overcome the limitations of VGG in weight initialization is Batch Normalization (BN) [21]. Batch normalization operates on mini-batches by normalizing the activation distribution at each layer of the network [22]. This technique helps maintain

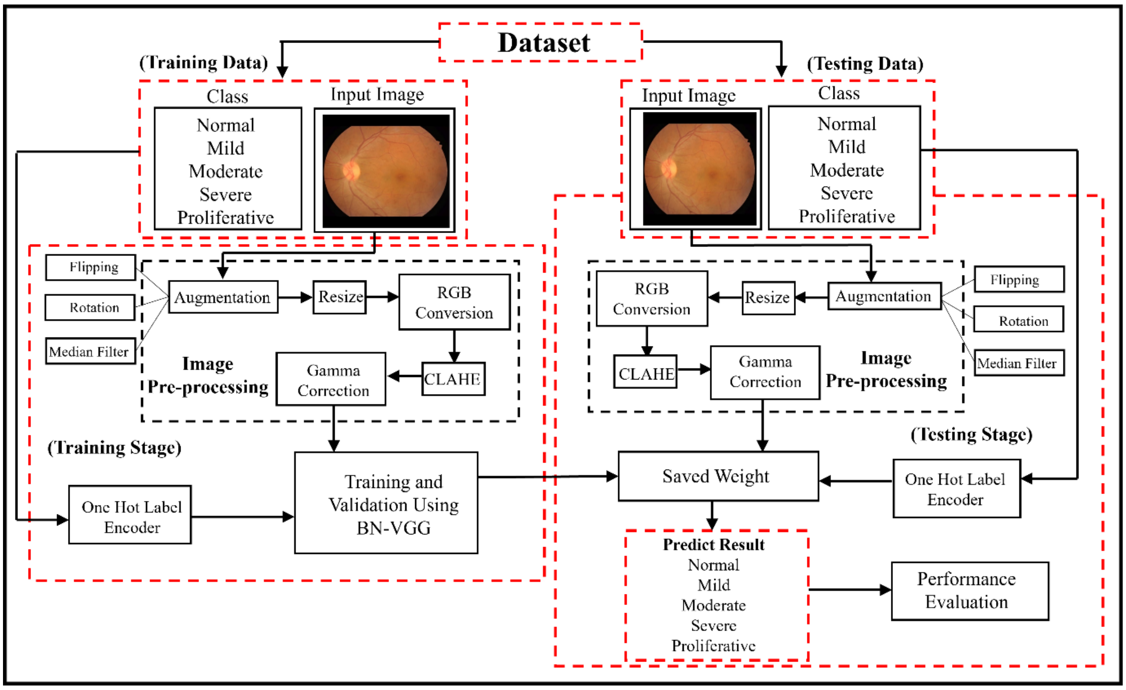
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

stable activation values, preventing them from becoming excessively large (exploding gradients) or too small (vanishing gradients). Batch normalization improves training stability, accelerates convergence, and enhances model generalization, particularly when training data are limited [23]. Slamet [24] applied batch normalization to a CNN for diabetic retinopathy classification, resulting in a 9% improvement in classification performance compared to a model without batch normalization. However, the study only considered four classes and did not apply batch normalization to the VGG architecture. Athira and Nair [25] applied batch normalization to VGG, which resulted in an accuracy of 86%, but this study performed diabetic retinopathy classification in 3 classes.

This study introduces a deep learning framework that incorporates gamma correction-based image enhancement with a Batch Normalization–Visual Geometry Group (BN-VGG) architecture for diabetic retinopathy classification, use retinal fundus images. Gamma correction is applied to improve brightness and reduce illumination variability, while BN-VGG enhances training stability, convergence speed, and generalization performance. The proposed method classifies diabetic retinopathy into 5 classifications: normal, mild, moderate, severe, and proliferative. The following summarizes the major contributions of this study: (1) the integration of gamma correction to improve retinal image quality prior to classification, (2) the development of a BN-VGG model to improve classification performance, and (3) a systematic evaluation of the combined effect of image enhancement and BN-VGG using PSNR and SSIM for image quality assessment, as well as accuracy, specificity, sensitivity, Cohen's Kappa, F1-score, and G-Mean for classification performance. The model's effectiveness is validated through experimental comparisons with baseline models, including VGG and BN-VGG, which do not include image enhancement. Data augmentation is applied to address data limitations and support robust training. This study aims to support more reliable and efficient early detection of diabetic retinopathy.

MATERIALS AND METHODS

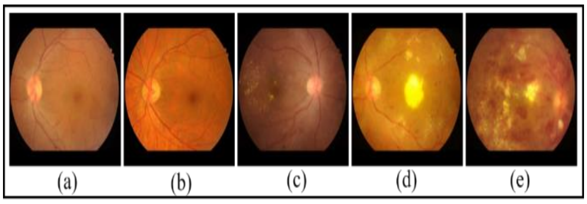
This study employs a BN-VGG architecture for diabetic retinopathy severity classification, with image quality enhanced using gamma correction as a preprocessing step. The overall proposed framework is demonstrated in Figure 1.



Source: (Research Results, 2026)
Figure 1. Proposed Framework for DR Severity Classification with Image Enhancement and BN-VGG Architecture

Data Description

This study uses the Indian Diabetic Retinopathy Image Dataset (IDRID), which is a publicly accessible database of retinal fundus images collected for the analysis of DR severity[26]. The dataset comprises 413 retinal images categorized into 5 classes different DR severity levels: 134 normal images, 20 mild images, 136 moderate images, 74 severe images, and 49 proliferative images. Figure 2 presents retinal fundus images obtained from the IDRID dataset.

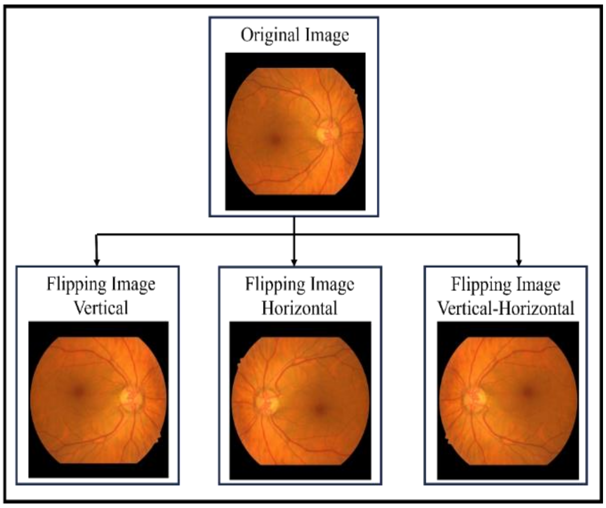


Source: (Research Results, 2026)

Figure 2. IDRID Images for Labels (a) Normal, (b) Mild, (c) Moderate, (d) Severe, and (e) Proliferative

Data Augmentation

Augmentation is performed to address the limitations of image data and enhance the learning capability of the model [5], [18], [19]. In this study, augmentation was performed using flipping, rotation, and median filtering techniques. The flipping technique is used to increase image variation without altering the features contained in the image [5]. The flipping was applied vertically, horizontally, and both vertically and horizontally. An example of the augmented images produced using vertical, horizontal, and combined vertical-horizontal flipping presented in Figure 3.



Source: (Research Results, 2026)

Figure 3. Example of Original Image and Flipping Augmentation Results View

Image Preprocessing

1. Resize
   Image resizing was performed to standardize input dimensions and ensure compatibility with the model [5]. All retinal fundus images have been resized to 256 × 256 pixels to provide a consistent input size while preserving relevant anatomical structures.
2. RGB Conversion
   The dataset contains retinal images that are represented in the blue, green, and red (BGR) channels. At this stage, the images are converted from BGR to red, green, and blue (RGB) channels. This conversion ensures that the image matches the model used and that the acquisition of the image channels is more accurate.
3. Contrast Limited Adaptive Histogram Equalization (CLAHE)
   CLAHE enhances the local contrast of image [27], [28], [29]. CLAHE applies histogram equalization to the values of each pixel in an image, revealing hidden features [29]. It works by setting a maximum limit on the height of the pixel histogram, known as the clip limit [28], [30]. The image is divided into several equally sized regions by the clip limit so that the specified distribution parameters are met by the histogram of the resulting image. The process of calculating the clip limit is described using Equation (1)[30].

$$\beta = \frac{d}{z}\left(1 + \frac{\alpha}{100}(x_{max} - 1)\right) \quad (1)$$

   where β denotes the clip limit, d represents the size of the region, z refers to the gray scale value, and α indicates the clip factor.
4. Gamma Correction
   The CLAHE image from the previous stage is corrected using gamma correction. The brightness levels of an image are adjusted using a process called gamma correction [29], [31], [32]. Gamma correction is a non-linear function. It serves as an intensity transformation function. It is defined using Equation (2) [29], [30].

$$I_{gam}(s,t) = I(s,t)^{\gamma} \quad (2)$$

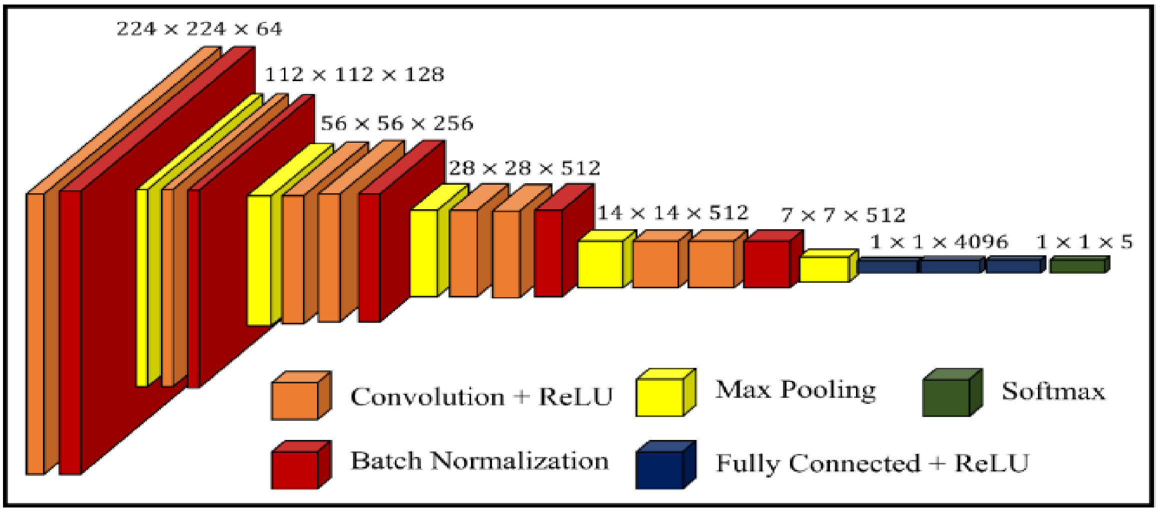
   for $s = 1,2,3,\dots,i$ and $t = 1,2,3,\dots,j$, where $I_{gam}(s,t)$ is the image pixel value at position (s,t), γ is the gamma value, i is the row number of pixel matrix, j is the column number of pixel matrix, γ is the gamma value, and $I_{gam}(s,t)$ is the result of gamma correction. A value of $\gamma > 1$ will

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

increase the contrast intensity, while a value of $\gamma < 1$ will decrease the contrast intensity.

Classification Using Batch Normalization-Visual Geometry Group (BN-VGG) Architecture

The classification stage of DR severity consists of two stages: training and testing. The retinal images resulting from image pre-processing are divided into two sets: 20% testing and 80% training. During training, the training data is split into training data and validation data. The training process aims to train the model to learn features of images. The training process begins with the initialization of parameters. These parameters include the epochs, batch size, and loss function. The model training uses the Batch Normalization-Visual Geometry Group (BN-VGG) architecture. The BN-VGG architecture combines the VGG architecture with batch normalization. In this architecture, batch normalization is added after each convolutional layer. Combination of the VGG architecture with batch normalization aims to facilitate the initialization of VGG weights, accelerate computation, improve convergence, and increase the classification accuracy. An illustration of the BN-VGG architectural structure presented in Figure 4.



Source: (Research Results, 2026)

Figure 4. Batch Normalization-Visual Geometry Group (BN-VGG) Architecture for DR Severity Classification on IDRID

Based on Figure 4, the image pixel matrix undergoes a process consisting of five convolutional blocks, each incorporating BN and max pooling. The process ends with fully connected and softmax layer. The convolution is designed to learn representations of features from the image. This layer contains a set of convolutional kernels used to extract the image's localized features [3]. The convolution is followed by the ReLU function. ReLU is an activation function that outputs zero if the input is negative [33]. The convolution result matrix is continued with the BN process. BN normalizes each layer of the network [34]. Max pooling (yellow block) is a dimension reduction operation of each normalization result. Max pooling uses 2 × 2 filter with stride 2. The fully connected layer and ReLU blocks (blue blocks) function as classifiers. The softmax block (green block) functions to convert the numeric output from the fully connected block into probabilities for each class [35], [36]. During training, the data were trained using the BN-VGG architecture with settings: 100 epochs, 6 batch size, 0.00001 learning rate, and softmax loss function.

The training process is carried out for the number of epochs specified during initialization. In the first epoch, the weights are initialized and subsequently updated based on validation data performance in each epoch. The best-performing weights are then used during the testing phase. In the testing phase, BN-VGG architecture model is used to make predictions on the test data. This stage evaluates the model's ability in classifying the severity of DR in retinal images.

Performance Evaluation

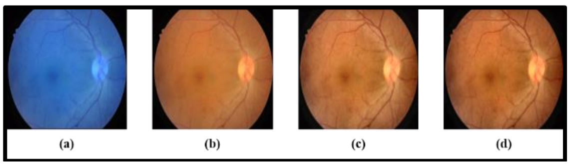
The effectiveness of the gamma correction method for enhancing quality of images was measured using Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) values. The performance of the proposed BN-VGG architecture in classifying the severity of DR in retinal images was measured using accuracy, specificity , sensitivity, Cohen's kappa, F1-score, and G-mean . These metrics were calculated using a confusion matrix, which is a table employed to illustrate the number of correct and incorrect predictions produced by a model [37], [38]. There are four main components: true positive, which represents correctly predicted positive labels; false positive, which represents negative labels that were incorrectly predicted as positive; true negative, which represents correctly predicted negative labels; and false negative, which represents positive labels that were incorrectly predicted as negative. [37], [38].

RESULTS AND DISCUSSION

Image Preprocessing

The image is augmented in the first stage to increase the amount of data and balance the labels on the data. The original image data consisted of 413 images consistin of 134 normal images, 20 mild images, 136 moderate images, 74 severe images, and 49 proliferative DR images. The images were enhanced using rotation, flip, and median filtering. The applied rotation technique consists of angles of 15°, 35°, 45°, 90°, and 270°, while the applied flip technique consists of horizontal, vertical, and vertical-horizontal flip. The application of the augmentation technique produces 12,372 data consisting of 2,496 normal labels, 2,400 mild labels, 2,496 moderate labels, 2,556 severe labels, and 2,424 proliferative DR labels. In the second step, the augmented image is resized so that the image has the same size. As a result of the resizing process, the image is displayed in the BGR (Blue-Green-Red) channel, which is then converted to the RGB channel. (Red-Green-Blue) channel.

The RGB images were first processed using CLAHE to enhance local contrast. CLAHE was implemented with clip limit of 1.3 and tile grid size of 4 × 4. The enhanced images were subsequently processed using gamma correction to normalize illumination and reduce contrast variability introduced during enhancement. A fixed gamma value of 1.2 was applied to flatten uneven contrast while preserving structural retinal features. This two-stage enhancement was designed to improve feature visibility prior to classification. The study's pre-processing steps and image enhancement results are displayed in Figure 5.



Source: (Research Results, 2026)

Figure 5. Results of The Preprocessing Stage on The Retinal Image (a) BGR Channel (b) RGB Channel Conversion (c) CLAHE (d) Gamma Correction

Based on Figure 5, the gamma correction process enhances image contrast, resulting in clearer and more distinguishable retinal features compared to the original RGB images. This improves highlighting lesion-related structures and supporting more effective classification.

The quality of the enhanced images is evaluated use SSIM and PSNR by comparing the original with pre-processed images. The average PSNR of the proposed preprocessing method is 30.85 dB, and the average SSIM is 0.87, indicating that the enhanced images show low distortion and structural similarity to the original images. These findings indicate that combining CLAHE with gamma correction effectively improves image contrast and reduces noise without degrading important visual information. The improved image quality is expected to contribute to better classification performance, as clearer feature representation can facilitate the learning process of model. To assess the effectiveness of the image enhancement method, the obtained PSNR and SSIM values are compared with those reported in previous studies on retinal image enhancement. The results of the comparative analysis are provided in Table 1.

Based on Table 1, the proposed image enhancement method that combines CLAHE and gamma correction shows better performance than existing approaches. The proposed method achieved the highest PSNR and SSIM, indicating an effective balance between noise reduction and structural preservation. These outcomes indicate that combining CLAHE with gamma correction is capable of enhancing image contrast while maintaining important structural details within the image.

Table 1. Comparison Result of PSNR and SSIM on This Study with Other Studies

No	Method	PSNR	SSIM
1	Convolutional Autoencoder [39]	29.1	0.78
2	RFormer [40]	27.42	0.84
3	Lab Color Space-Based Enhancement [41]	29.77	0.63
4	Cycle CBAM [42]	24.7	0.81
5	Proposed Method	**30.85**	**0.87**

Source: (Research Results, 2026)

Chilukuri et al. [39] and Priyadharsini et al. [41] produced relatively high PSNR values of 29.1 dB and 29.77 dB, respectively, but lower SSIM values, indicating that structural information was not preserved as effectively. Deng et al. [40] achieved a relatively high SSIM value of 0.84, suggesting good structural preservation, although its PSNR remained lower than the proposed

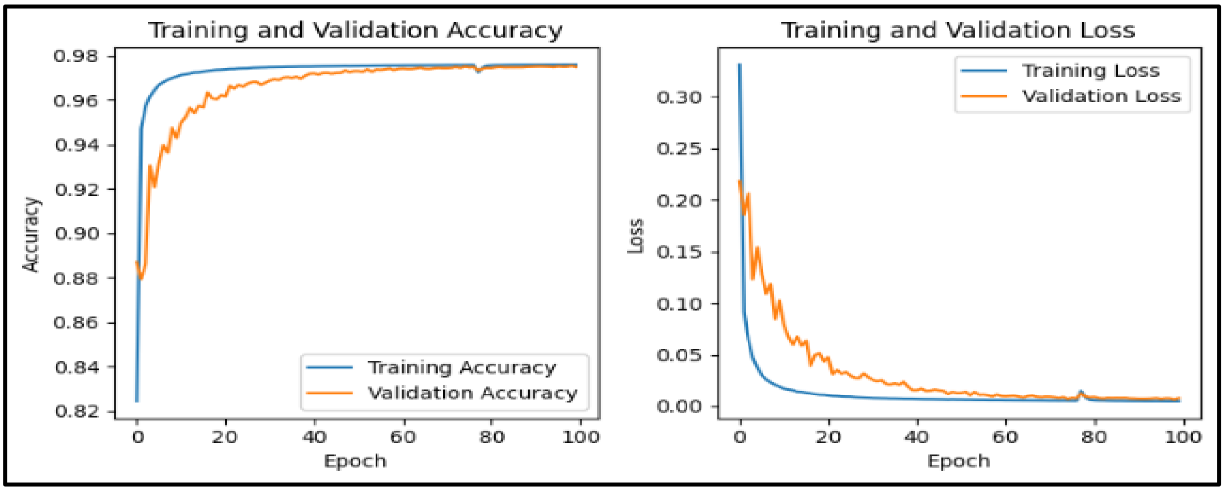
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

method. Wan et al. [42] obtained the lowest PSNR among methods with SSIM evaluation, indicating higher reconstruction distortion. Overall, the proposed preprocessing approach provides more balanced, stable, and reliable image quality improvement compared to existing methods. The use of RGB space also helps preserve important visual information, contributing to better overall image representation.

Classification Using Batch Normalization-Visual Geometry Group (BN-VGG) Architecture

The classification process consists of two steps: training and testing. The dataset is split into training data and testing data at the start of the training stage. The training data comprises 80% of the dataset, or 9,897 images, while the testing data is 20% of the whole data set or 2,475 images. The process was conducted using the Jupyter Notebook application on a 64-bit Windows 11 computer with 16GB of RAM, 1TB of storage, an Intel Core i5-12400 2.50GHz, and NVIDIA GeForce RTX 3060. The labels used in this study are normal, mild, moderate, severe, and proliferative DR. On the training stage, the training data is also splited into two parts: 80% or 7,917 images for training and 20% or 1,980 images for validation. Data was augmented only for the training dataset to avoid leakage during validation and testing. Although cross-validation could provide additional robustness, this study focuses on a controlled, comparative evaluation with a consistent data split to analyze the effect of the proposed enhancement and architecture. Validation data is used to assess the model's performance on unseen data during training, detect overfitting, and determine if the model can effectively perform on different data. During training, accuracy and loss values are calculated using training data and validation data. The efficacy of the model's classification is measured by accuracy, while loss quantifies the difference between the predicted and actual labels. Figure 6 shows accuracy versus loss graphs during training process.



Source: (Research Results, 2026)

Figure 6. Graphs of (a) Accuracy and (b) Loss During Training of DR Severity Classification Using The BN-VGG Architecture

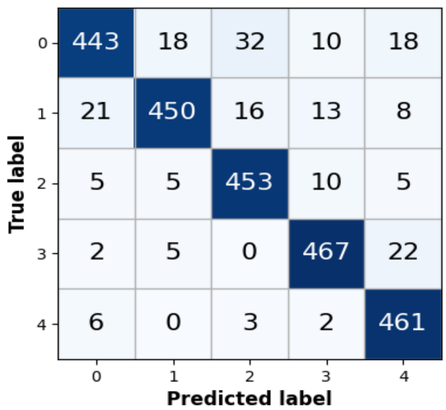
Based on Figure 6(a), the accuracy value is above 0.96 at the training stage. For both the training data and validation data, this value keeps rising and then steadies itself at the 50th epoch. However, accuracy value drops slightly at the 80th epoch before increasing again until the end. Figure 6(b) shows the decrease in loss value for both training and validation data at the training stage, with a steady decline towards a value below 0.05. Additionally, both the training and validation losses decrease and stabilize consistently during the training process. This behavior indicates stable learning and suggests that, despite the limited dataset size, the proposed model does not suffer from severe overfitting.

During the training stage, the best weight obtained will be saved for prediction in the testing process. The testing data is 2,475 images. This data is used to evaluate the performance of the trained model in image classification. Its performance will be evaluated by measuring accuracy, specificity, sensitivity, F1 score, and Cohen's kappa for each label. The performance evaluation results have very good values, especially for the severe and proliferative labels, which have values above 0.92

for each performance. The specificity obtained for each label has value above 0.96. This shows the model can identify data that isn't included in a specific label. The sensitivity and Cohen's kappa for the normal label are only 0.85 and 0.86, but the accuracy, F1-score, specificity, and G-Mean values are above 0.89. The proliferative label has the highest sensitivity value, but its specificity value is still below the mild label. Overall, each label has good performance evaluation results. This proves the architecture's effectiveness and accuracy in identifying the severity of DR disease.

Performance Evaluation

At the classification stage, the training process has accuracy and sensitivity value above 0.95, while specificity, F1-score, and Cohen's kappa above 0.9, and loss is below 0.05. At the testing stage, the classification stage has an average accuracy, specificity, and G-Mean value above 0.96, sensitivity and F1-score above 0.91, and Cohen's kappa of 0.9. This indicates that the proposed classification model can effectively classify diabetic retinopathy into its respective severity categories. To further analyze the classification performance and misclassification patterns across different classes, the confusion matrix of the model is presented in Figure 7.



Source: (Research Results, 2026)

Figure 7. Confusion Matrix of The Proposed Model on The Retinal Image Test Set

As shown in Figure 7, the following corresponds to the labels: 0 is normal, 1 is mild, 2 is moderate, 3 is severe, and 4 is proliferative. The normal class achieved 443 correct predictions, with a small number of samples misclassified, mainly as moderate. The mild class obtained 450 correct predictions, with only a few misclassifications distributed across the other classes. The moderate class also demonstrated strong performance, with 453 correctly predicted images and limited errors in other categories. Similarly, the severe class achieved 467 correct predictions, although several samples were misclassified as proliferative. The proliferative class recorded 461 correct predictions with relatively few misclassifications.

Most misclassifications occurred between severity levels, for example, normal and moderate or severe and proliferative. This pattern suggests that the model may have difficulty distinguishing cases with subtle or overlapping retinal features, which is common in diabetic retinopathy grading due to the gradual progression of lesions across severity levels. Nevertheless, the small quantity of off-diagonal values suggests that the proposed model maintains strong discriminative capability across different stages of the disease.

In addition to the confusion matrix analysis, several cases of misclassification were observed during the testing phase. For example, some images belonging to the Mild class were predicted as Normal, which may occur because early diabetic retinopathy lesions, such as microaneurysms, are often very small and difficult to distinguish from normal retinal structures. Similarly, a few severe cases were misclassified as proliferative, likely due to similarities in retinal abnormalities, such as extensive hemorrhages and vascular changes, that appears in the disease's advanced stages. These findings suggest that misclassifications occur primarily when retinal features are subtle or pathological characteristics overlap between adjacent severity levels.

To validate the contribution of each element in the proposed workflow, an experimental study was conducted by comparing three configurations: (1) baseline VGG without enhancement or batch normalization, (2) BN-VGG without image enhancement, and (3) the proposed model. The summary of the results is demonstrated in Table 2.

Table 2. Experimental Study of Image Enhancement and BN-VGG Architecture for Diabetic Retinopathy Classification

No	Method	Acc	Sen	Spe	F1-Score	Cohen's Kappa	G-Mean
1	Baseline VGG	0.83	0.83	0.94	0.83	0.78	0.82
2	BN-VGG without enhancement image	0.85	0.85	0.95	0.85	0.8	0.84
3	BN-VGG+Gamma Correction (Proposed Method)	**0.97**	**0.92**	**0.98**	**0.92**	**0.9**	**0.97**

Source: (Research Results, 2026)

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

Based on Table 2, the model performance gradually improved from the baseline VGG to BN-VGG without image enhancement, and further to BN-VGG with gamma correction, which represents the proposed approach. The baseline VGG achieved an accuracy of 0.83 with Cohen's kappa of 0.78, indicating moderate agreement in classification performance. The introduction of batch normalization slightly improved the results, increasing accuracy to 0.85 and Cohen's kappa to 0.80. This improvement suggests that batch normalization contributes to more stable training and improved feature learning.

The most significant performance improvement was observed when gamma correction was integrated with the BN-VGG architecture. The proposed method achieved strong classification performance, with sensitivity, F1-score, and Cohen's kappa above 0.89, while accuracy, specificity, and G-mean were above 0.96. These results indicate that the proposed enhancement improves feature visibility and enables the capture of relevant retinal patterns. Overall, the combination of batch normalization and gamma correction improves classification performance compared to the baseline and BN-VGG without enhancement.

The results of this study's classification are examined alongside those of previous studies. The comparative results are demonstrated in Table 3. Based on Table 3, this study obtained the highest accuracy, specificity, G-mean, Cohen's kappa, and F1-score compared to other studies. Although Abdelmaksoed et al. [43] reported a higher sensitivity, their model shows lower specificity and does not include F1-score evaluation, indicating a potential imbalance in classification performance.

Table 3. Performance Comparison of The Proposed Model with Other DR Classification Methods

No	Method	Acc	Sen	Spe	F1-Score	Cohen's Kappa	G-Mean
1	Multi-Scale Attention-ResNet Gradient Boosting [44]	0.94	0.91	-	0.91	-	-
2	EfficientNet-B0 [45]	0.86	**-**	-	-	-	-
3	E-DenseNet BC-121 [43]	0.91	**0.96**	0.69	-	0.88	0.81
4	E-DenseNet BC-201-ImageNet [43]	0.62	0.61	0.55	-	0.48	0.58
5	BN-VGG (Proposed Method)	**0.97**	0.92	**0.98**	**0.92**	0.9	**0.97**

Source: (Research Results, 2026)

In contrast, the proposed method achieves a more balanced performance, with high sensitivity and specificity, demonstrating its effectiveness in both detecting DR cases and correctly identifying non-DR samples. The study by Srinivasan et al. [44] had a high accuracy value compared to Abdelmaksoud et al. [43], [45], but did not report specificity or Cohen's kappa. Based on the outcome of the performance evaluations, the proposed BN-VGG architecture showed very good ability in classifying the severity of diabetic retinopathy accurately. Although the sensitivity value obtained was slightly lower compared to the study by Barakat et al. [43], it remains above 0.9, indicating strong capability in detecting DR cases. Moreover, the high specificity value of 0.98 demonstrates that the model can accurately identify non-DR samples, reducing the likelihood of false positives. The Cohen's Kappa of 0.90 further indicates strong agreement between predicted and actual labels. This confirms the robustness of the proposed BN-VGG architecture for classifying DR severity.

The experimental results demonstrate that the proposed enhancement-classification pipeline improves both visual quality and diagnostic performance. The integration of gamma correction with BN-VGG contributes to stable convergence and enhanced feature extraction, particularly for intermediate severity levels where lesion patterns are subtle. Compared with previous studies, the proposed approach achieves competitive performance while maintaining architectural simplicity. These findings suggest that appropriate preprocessing can significantly influence deep learning-based DR severity classification. Additionally, the suggested method has the capacity to facilitate the implementation of AI-powered screening tools for the timely identification of diabetic retinopathy, particularly in large-scale clinical settings where automated analysis can assist ophthalmologists in decision-making. The study had some weaknesses. Initially, the model was trained and assessed using a modest dataset. However, the generalization ability of the proposed approach may be limited when applied to larger, more diverse clinical datasets. Second, this analysis was based on a single benchmark dataset. Therefore, application to images acquired from different devices or clinical environments may produce different results. In addition, this study focuses on a specific enhancement strategy and architecture configuration, and further investigation using different enhancement techniques or deep learning architectures may provide additional insights.

CONCLUSION

This study proposes a method for classifying diabetic retinopathy (DR) severity using a deep learning-based approach by integrating gamma correction-based image enhancement with a Batch Normalization–Visual Geometry Group (BN-VGG) architecture. The results demonstrate that the suggested method improves image quality while achieving high and balanced classification performance across five DR severity levels.

This study's most significant finding is the incorporation of an effective yet simple image enhancement technique into the BN-VGG model. This integration improves feature representation and produces reliable classification results. However, the study's use of a relatively small dataset and evaluation on a single benchmark dataset may affect the model's ability to generalize. Future research will center on validating the model using larger, more diverse datasets and exploring additional enhancement techniques to improve robustness and performance.

REFERENCE

[1] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey," *IEEE Access*, vol. 10, pp. 28642–28655, 2022, doi: 10.1109/ACCESS.2022.3157632.

[2] Z. Khan *et al.*, "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408–61416, 2021, doi: 10.1109/ACCESS.2021.3074422.

[3] A. Bilal, G. Sun, and S. Mazhar, "Survey on recent developments in automatic detection of diabetic retinopathy," *J. Fr. Ophtalmol.*, vol. 44, no. 3, pp. 420–440, 2021, doi: 10.1016/j.jfo.2020.08.009.

[4] N. A. El-Hag *et al.*, "Classification of retinal images based on convolutional neural network.," *Microsc. Res. Tech.*, vol. 84, no. 3, pp. 394–414, Mar. 2021, doi: 10.1002/jemt.23596.

[5] A. Desiani *et al.*, "Multi-Stage CNN: U-Net and Xcep-Dense of Glaucoma Detection in Retinal Images," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 5, no. 4, pp. 211–222, 2023, doi: 10.35882/jeeemi.v5i4.314.

[6] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, pp. 1–22, 2021, doi: 10.3390/s21113704.

[7] N. S. Rekhi, J. S. Sidhu, and A. Arora, "Gamma Correction for Brightness Preservation in Natural Images," *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. 2791–2807, 2023, doi: 10.32604/csse.2023.026976.

[8] A. Desiani, D. P. Rini, and L. I. Kesuma, "DDUSeg-Net as a Design of Convolutional Neural Network Architecture for Semantic Segmentation in Cervical Cancer.," *Int. J. Intell. Eng. Syst.*, vol. 18, no. 1, 2025, doi: 10.22266/ijies2025.0229.70.

[9] A. Desiani, Erwin, B. Suprihatin, F. Efriliyanti, M. Arhami, and E. Setyaningsih, "VG-DropDNet a Robust Architecture for Blood Vessels Segmentation on Retinal Image," *IEEE Access*, vol. 10, pp. 92067–92083, 2022, doi: 10.1109/ACCESS.2022.3202890.

[10] S. Soni, P. Singh, and A. Waoo, "Review of Gamma Correction Techniques in Digital Imaging," *ShodhKosh J. Vis. Perform. Arts*, vol. 5, May 2024, doi: 10.29121/shodhkosh.v5.i5.2024.1902.

[11] J. Xue, Y. Wang, Z. Li, and H. Zhang, "Classification of Diabetic Retinopathy Based on Efficient Computational Modeling," Applied Sciences, vol. 14, no. 23, p. 11327, 2024, doi: 10.3390/app142311327.

[12] T. Aziz, C. Charoenlarpnopparut, and S. Mahapakulchai, "Deep learning-based hemorrhage detection for diabetic retinopathy screening," Scientific Reports, vol. 13, no. 1, p. 1479, 2023, doi: 10.1038/s41598-023-28680-3.

[13] D. Mane, R. Ashtagi, R. Suryawanshi, A. N. Kaulage, and A. N. Hedaoo, "Diabetic Retinopathy Recognition and Classification Using Transfer Learning Deep Neural Networks," *Int. Inf. Eng. Technol. Assoc.*, vol. 41, no. 5, pp. 2683–2691, 2024, doi: 10.18280/ts.410541.

[14] Z. Yang, "Classification of picture art style based on VGGNET," *J. Phys. Conf. Ser.*, vol. 1774, no. 1, pp. 0–9, 2021, doi: 10.1088/1742-6596/1774/1/012043.

[15] D. Abhishek and P. Jatin, "Automated detection of Diabetic Retinopathy using VGG-16 architecture," *Int. Res. J. Eng. Technol.*, vol. 8, no. 3, pp. 2936–2940, 2021.

[16] A. T. Nair, M. L. Anitha, and M. N. Arun Kumar, "Disease Grading of Diabetic Retinopathy using Deep Learning Techniques," *Proc. - 6th Int. Conf. Comput. Methodol. Commun. ICCMC 2022*, no. Iccmc, pp. 1019–1024, 2022, doi: 10.1109/ICCMC53470.2022.9754113.

[17] E. M. F. El Houby, "Using transfer learning for diabetic retinopathy stage classification," *Appl. Comput. Informatics*, 2021, doi: 10.1108/ACI-

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8094

07-2021-0191.
[18] J. Aftab, M. A. Khan, S. Arshad, S. ur Rehman, D. A. AlHammadi, and Y. Nam, "Artificial Intelligence Based Classification and Prediction of Medical Imaging using A Novel Framework of Inverted and Self-Attention Deep Neural Network Architecture," *Sci. Rep.*, vol. 15, no. 8724, pp. 1–26, 2025, doi: 10.1038/s41598-025-93718-7.
[19] A. S. Farhan, M. Khalid, and U. Manzoor, "Combined Oriented Data Augmentation Method for Brain MRI Images," *IEEE Access*, vol. 13, pp. 9981–9994, 2025, doi: 10.1109/ACCESS.2025.3526684.
[20] H. Yang, J. Ni, J. Gao, Z. Han, and T. Luan, "A novel method for peanut variety identification and classification by Improved VGG16," *Sci. Rep.*, vol. 11, no. 1, pp. 1–17, 2021, doi: 10.1038/s41598-021-95240-y.
[21] A. Desiani, Erwin, B. Suprihatin, Y. Wahyudi, E. S. Cahyono, and M. Arhami, "BVU-Net : A U-Net Modification by VGG-Batch Normalization for Retinal Blood Vessel Segmentation," *Int. J. Intell. Eng. Syst.*, vol. 15, no. 6, pp. 303–314, 2022, doi: 10.22266/ijies2022.1231.29.
[22] L. Huang, J. Qin, Y. Zhou, F. Zhu, L. Liu, and L. Shao, "Normalization techniques in training dnns: Methodology, analysis and application," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10173–10196, 2023, doi: 10.1109/TPAMI.2023.3250241.
[23] E. Lobacheva, M. Kodryan, N. Chirkova, A. Malinin, and D. P. Vetrov, "On the periodic behavior of neural network training with batch normalization and weight decay," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 21545–21556, 2021, doi: 10.48550/arXiv.2106.15739.
[24] I. Slamet, "Retinopathy Classification using Convolutional Neural Network Method with Adaptive Momentum Optimization and Applied Batch Normalization," *J. Appl. Data Sci.*, vol. 5, no. 3, pp. 1123–1133, 2024, doi: 10.47738/jads.v5i3.309.
[25] T. R. Athira and J. J. Nair, "Diabetic Retinopathy Grading from Color Fundus Images: An Autotuned Deep Learning Approach," *Procedia Comput. Sci.*, vol. 218, no. 2022, pp. 1055–1066, 2022, doi: 10.1016/j.procs.2023.01.085.
[26] P. Porwal, S. Pachade, M. Kokare, G. Deshmukh, J. Son, W. Bae, L. Liu, J. Wang, X. Liu, L. Gao, T. Wu, J. Xiao, F. Wang, B. Yin, Y. Wang, G. Danala, L. He, Y. H. Choi, Y. C. Lee, S.-H. Jung, et al., "IDRiD: Diabetic Retinopathy—Segmentation and Grading Challenge," Medical Image Analysis, vol. 59, p. 101561, 2020, doi: 10.1016/j.media.2019.101561.
[27] M. Triantafyllou *et al.*, "The Effect of Ultrasound Image Pre-Processing on Radiomics Feature Quality : A Study on Shoulder Ultrasound," *J. Imaging Informatics Med.*, 2025, doi: 10.1007/s10278-025-01421-w.
[28] V. Madhavi and P. L. S. Kumari, "A Qualitative Approach for Enhancing Fundus Images with Novel CLAHE Methods," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 20102–20107, 2025, doi: 10.48084/etasr.9525.
[29] M. N. Huda, A. Musdholifah, and A. Z. K. Frisky, "Optimization of Plantar Foot Thermogram for Diabetic Foot Ulceration Early Detection : An Image Enhancement Approach," *SINTECHCOM Sci. Technol. Commun. J.*, vol. 5, no. 2, pp. 49–66, 2025, doi: 10.59190/stc.v5i2.273.
[30] C. P. Parmo, L. I. Kesuma, and D. Geovani, "The Combination of Black Hat Transform and U-Net in Image Enhancement and Blood Vessel Segmentation in Retinal Images," *Comput. Eng. Appl.*, vol. 12, no. 3, pp. 129–145, 2023, doi: 10.18495/comengapp.v12i3.452.
[31] M. Naufal, H. Al Azies, and R. M. Brilianto, "Enhanced Brain Tumor Classification Through Gamma Correction in Deep Learning," *Sist. J. Sist. Inf.*, vol. 13, no. 6, pp. 2348–2358, 2024, doi: 10.32520/stmsi.v13i6.4474.
[32] M. N. Noor, F. Haneef, I. Ashraf, and M. Masud, "Enhanced Skin Disease Classification via Dataset Refinement and Attention-Based Vision Approach," *Bioengineering*, vol. 12, no. 275, pp. 1–17, 2025, doi: 10.3390/bioengineering12030275.
[33] P. Radočaj, D. Radočaj, and G. Martinović, "Optimal Optimizing Convolutional Neural Network Architectures with Optimal Activation Functions for Pediatric Pneumonia Diagnosis Using Chest X-rays," *Big Data Cogn. Comput.*, vol. 9, no. 25, pp. 1–17, 2025, doi: 10.3390/bdcc9020025.
[34] D. F. Brianna, L. I. Kesuma, D. Geovani, and P. Sari, "Combination of Image Enhancement and Double U-Net Architecture for Liver Segmentation in CT-Scan Images," *J. Electron. Electromed. Eng. Med. Informatics*, vol. 7, no. 1, pp. 208–219, 2025, doi: 10.35882/jeeemi.v7i1.582.
[35] M. N. Hasan, M. E. R. Pial, S. Das, N. Siddique, and H. Wang, "DIA-VXNET: A Framework for Automated Diabetic Eye Disease Detection using Transfer Learning with Feature Fusion Network," *Biomed. Signal Process. Control*, vol. 100, no. 106907, pp. 1–23, 2025, doi:

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

10.1016/j.bspc.2024.106907.

[36] S. Afrifa, V. Varadarajan, P. Appiahene, T. Zhang, D. Gyamfi, and R. M. O. M. Gyening, “Deep Neural Networks for Skin Cancer Classification: Analysis of Melanoma Cancer Data,” *J. Adv. Inf. Technol.*, vol. 16, no. 1, pp. 1–11, 2025, doi: 10.12720/jait.16.1.1-11.

[37] M. Abidin *et al.*, “Classification of Heart (Cardiovascular) Disease using the SVM Method,” *Indones. J. Mod. Sci. Technol. Mod.*, vol. 1, no. 1, pp. 9–15, 2025, doi: 10.64021/ijmst.1.1.9-15.2025.

[38] R. Wijaya and G. Kosala, “Stress Detection Through Wearable Sensors : A Convolutional Neural Network-Based Approach Using Heart Rate and Step Data,” *Int. J. Electr. Comput. Eng.*, vol. 15, no. 2, pp. 1880–1888, 2025, doi: 10.11591/ijece.v15i2.pp1880-1888.

[39] R. Chilukuri, P. Praveen, R. K. Gatla, and R. A. Almenweer, “Quantum Denoising Autoencoder Improves Retinal Fundus Image Quality for Early Diabetic Retinopathy Screening,” *Sci. Rep.*, vol. 16, no. 5970, pp. 1–17, 2026, doi: 10.1038/s41598-026-35540-3.

[40] Z. Deng, Y. Cai, L. Chen, Z. Gong, Q. Bao, X. Yao, D. Fang, W. Yang, S. Zhang, and L. Ma, “RFormer: Transformer-Based Generative Adversarial Network for Real Fundus Image Restoration on a New Clinical Benchmark,” IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 9, pp. 4645–4655, Sep. 2022, doi: 10.1109/JBHI.2022.3187103.

[41] P. C and J. K. R, “Retinal image enhancement based on color dominance of image,” *Sci. Rep.*, vol. 13, no. 1, pp. 1–14, 2023, doi: 10.1038/s41598-023-34212-w.

[42] C. Wan *et al.*, “Retinal Image Enhancement Using Cycle-Constraint Adversarial Network,” *Front. Med.*, vol. 8, no. January, pp. 1–16, 2022, doi: 10.3389/fmed.2021.793726.

[43] E. AbdelMaksoud, S. Barakat, and M. Elmogy, “A Computer-Aided Diagnosis System for Detecting Various Diabetic Retinopathy Grades Based on A Hybrid Deep Learning Technique,” *Med. Biol. Eng. Comput.*, vol. 60, no. 7, pp. 2015–2038, 2022, doi: 10.1007/s11517-022-02564-6.

[44] V. Srinivasan and V. Rajagopal, “Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification,” *Int. J. Intell. Eng. Syst.*, vol. 15, no. 4, pp. 489–498, 2022, doi: 10.22266/ijies2022.0831.44.

[45] E. Abdelmaksoud, S. Barakat, and M. Elmogy, “Diabetic Retinopathy Grading System Based On Transfer Learning,” *Int. J. Adv. Comput. Res.*, vol. 11, no. 52, pp. 1–12, 2021, doi: 10.48550/arXiv.2012.12515.