The object of the study is the accuracy of diabetic retinopathy diagnosis based on retinal images. This study investigates convolutional neural network (CNN) models for the automatic detection of diabetic retinopathy (DR) from retinal images. The main problem lies in the insufficient effectiveness of basic CNN models in recognizing DR stages on fundus images. The core problem addressed is the suboptimal performance of baseline CNNs in identifying DR stages from medical imagery. To solve this, two CNN architectures were thoroughly evaluated: a baseline model and an enhanced model integrating advanced preprocessing techniques such as image resizing (256 × 256 and 512 × 512), the image normalization, and data augmentation methods. The enhanced model outperformed the original, achieving a validation accuracy of 91% compared to 88% for the baseline, and demonstrating reduced loss during both training and validation. This improvement is attributed to the optimized input image quality and increased variability in the training set, which enhanced the model's ability to generalize and avoid overfitting. A distinctive feature of the results lies in the synergy between preprocessing and CNN architecture, which enabled significantly improved classification performance even under hardware constraints. These limitations suggest that further gains are possible with extended computational resources and access to larger datasets. The practical applicability of the findings is evident in the potential deployment of such models in clinical screening systems to support early and accurate DR diagnosis. The models were trained on a proprietary dataset of expert-labeled, high-resolution retinal images, similar in format to EyePACS and APTOS, though not publicly available due to ethical considerations

Keywords: diabetic retinopathy, fundus images, image preprocessing, contrast enhancement, data augmentation

UDC 004.93
DOI: 10.15587/1729-4061.2025.335570

DEVELOPMENT OF AN IMAGE QUALITY ENHANCEMENT APPROACH FOR DIABETIC RETINOPATHY DIAGNOSIS

Saya Sapakova
Corresponding author
Candidate of Physical and Mathematical Sciences, Associate Professor*
E-mail: s.sapakova@iitu.edu.kz
Nurmaganbet Yesmukhamedov
PhD Candidate*
Askar Sapakov
Candidate of Technical Sciences, Assistant Professor
Department of Electric Power Engineering
ALT University
Shevchenko str., 97, Almaty,
Republic of Kazakhstan, 050012
*Department of Computer Engineering
International IT University
Manasa str., 34/1, Almaty,
Republic of Kazakhstan, 050060

Received 04.06.2025
Received in revised form 23.07.2025
Accepted date 18.08.2025
Published date 29.08.2025

How to Cite: *Sapakova, S., Yesmukhamedov, N., Sapakov, A. (2025). Development of an image quality enhancement approach for diabetic retinopathy diagnosis. Eastern-European Journal of Enterprise Technologies, 4 (9 (136)), 79–88. https://doi.org/10.15587/1729-4061.2025.335570*

1. Introduction

Diabetic retinopathy (DR) remains one of the most severe complications of diabetes mellitus and is the leading cause of preventable vision loss among working-age adults worldwide. According to the International Diabetes Federation, the global number of people with diabetes is steadily increasing, which inevitably leads to a rise in DR cases. Early diagnosis and accurate classification of DR stages are crucial to initiate timely treatment and reduce the risk of blindness. Traditional diagnostic methods based on manual analysis of fundus images are time-consuming, prone to inter-observer variability, and not scalable to large populations. In this context, artificial intelligence (AI), and in particular deep learning and convolutional neural networks (CNNs), have demonstrated promising results in the automated classification of DR. Research in this field is rapidly evolving and continues to draw attention from the global scientific community. Despite the progress, several fundamental challenges remain unresolved. One of the key scientific issues is the dependence of model accuracy on image quality. In real clinical scenarios, retinal images often suffer from poor resolution, uneven illumination, and various types of noise. These factors significantly impair the generalizability of neural network models. Improving the quality of input data before classification – through techniques such as image enhancement, denoising, and normalization – is a critical direction in modern research aimed at improving the robustness and diagnostic value of AI models. Recent studies show that integrating preprocessing methods with optimized neural network architectures can significantly improve diagnostic performance. However, there is still no unified approach or consensus in the research community regarding which preprocessing strategies are most effective in real-world clinical settings. Moreover, most publicly available datasets are limited in diversity, which restricts the transferability of proposed models to heterogeneous populations. Therefore, research on the development and integration of image enhancement techniques in deep learning-based diabetic retinopathy diagnostics remains a highly relevant and practically significant scientific topic. This direction directly contributes to the development of scalable, accurate, and accessible AI-driven tools in ophthalmology, which are essential for reducing the global burden of vision impairment caused by diabetes.

2. Literature review and problem statement

In the work [1], the problem of automated diabetic retinopathy diagnosis using image processing and deep learning

is studied; however, important aspects such as transparent description of preprocessing and training, dataset details, external validation, and comparison with advanced architectures remain unexplored, which can be explained by objective limitations of data access and methodological difficulties related to class imbalance and clinical validation. In [2], the problem of automatic selection of the optimal zone for retinal laser coagulation is studied; however, unexplored aspects include detailed image processing methods, model training specifics, external clinical validation, and comparison with existing approaches, which can be explained by limited annotated data, difficulties in standardizing expert labeling, and practical constraints of clinical testing. In the work [3], the problem of automatic diabetic retinopathy detection using morphological analysis and a clinical application prototype is studied; however, details of image processing algorithms, external validation, and comparison with existing methods remain unexplored, mainly due to limited annotated data, difficulties in standardizing expert labeling, and practical constraints of clinical integration. In [4], specialized filtering and image enhancement techniques for diabetic retinopathy analysis were proposed, including min-max filtering for noise suppression, pixel range enhancement to highlight pathological regions, and convolution with custom kernels for lesion detection. These methods significantly improved the clarity and diagnostic usability of fundus images. However, the study was limited to laboratory testing and did not include external clinical validation or comparison with deep neural networks. The identified gaps stem from objective dataset limitations and mathematical difficulties in integrating filtering methods with modern deep learning architectures.

In [5], the problem of enhancing retinal fundus images for diabetic retinopathy (DR) diagnosis was studied through contrast enhancement and optimization, with CLAHE proving most effective under poor lighting. However, the work does not address how these enhancements influence downstream classification accuracy or clinical decision-making. This gap is partly due to methodological constraints, as linking low-level enhancement to high-level diagnostic performance requires large annotated datasets and controlled validation, which were not available.

The authors [6] examined CNN architectures (AlexNet, VGG19, EfficientNetV2B0) for multi-class classification of cataract and glaucoma, finding EfficientNetV2B0 superior when paired with Grad-CAM interpretability. Yet, the study leaves unexplored the impact of preprocessing pipelines and broader model comparisons with state-of-the-art transformers. These omissions can be explained by both objective constraints of computational resources and methodological complexity in ensuring fair benchmarking across architectures. In [7], several denoising and enhancement filters – including Gaussian-bilateral, Gabor, Haar wavelet, and guided image filtering were compared for DR lesion visibility, with guided filtering achieving the best visual clarity. Nonetheless, the effect of such enhancement on classification accuracy remains untested, limiting its clinical value. This gap is rooted in methodological difficulties of integrating enhancement into end-to-end deep learning pipelines and in the lack of labeled datasets linking enhancement to diagnostic outcomes.

The problem of improving early DR detection was addressed by proposing pixel color amplification integrated with EfficientNetV2, which enhanced sensitivity to subtle pathological changes in [8]. However, the study does not explore generalization across diverse datasets or robustness to variations in imaging devices. The reason lies in objective limitations of available datasets and mathematical challenges of ensuring color normalization across heterogeneous imaging conditions.

The study in [9] compared morphological operations with region of interest (ROI) extraction to improve fundus image interpretability for DR screening. While effective for feature highlighting, the study did not evaluate how these methods integrate with deep learning classifiers or compare to advanced enhancement techniques. These omissions are explained by methodological constraints, as morphological techniques are simpler to implement and validate, while their integration with modern architectures requires more sophisticated experimental setups and annotated clinical data.

A collaborative learning framework was introduced in [10] to enhance low-quality fundus images while maintaining lesion structures, which significantly improved DR grading accuracy. Nevertheless, the study did not address scalability and robustness across different datasets, likely due to computational demands of the framework and the absence of external validation. Further refinements were suggested in [11], where histogram plotting, green channel extraction, PCA, and Canny filtering were applied to assist in microaneurysm and exudate detection. Despite the detailed preprocessing steps, the lack of integration with advanced deep learning architectures and limited dataset validation remain gaps, probably stemming from methodological constraints and difficulties in standardizing preprocessing pipelines.

Data augmentation methods for segmentation were explored in [12, 13], where Distributive and Generative Adversarial Networks (D-GAN, G-GAN) were applied for blood vessel extraction with the help of median filtering and background subtraction. These works contributed to improving vessel visibility, but they remain vulnerable to noise and variability in image quality. The main reason for this gap lies in mathematical and computational challenges associated with stabilizing GAN training and ensuring robustness under diverse acquisition conditions.

A different approach was presented in [14], where CLAHE was combined with GANs to increase classification accuracy from 95.4% to 98%. Similarly, [15, 16] employed enhanced CLAHE and hybrid filtering to mitigate issues of uneven illumination and low contrast, demonstrating noticeable improvements in lesion visibility. However, these studies lack a systematic evaluation against alternative preprocessing methods and fail to address real-time applicability, which may be explained by the focus on experimental performance rather than clinical translation. Work [17] instead emphasized CNN optimization without explicit preprocessing, achieving 91% accuracy. Yet, the absence of enhancement makes it difficult to assess robustness in poor-quality images, which suggests methodological limitations in designing models without preprocessing integration.

Subsequent research focused on filter-based methods. Study [18] proposed hybrid image filters and enhanced CLAHE to tackle low contrast and blurring, while [19] applied CLAHE with morphological transformations to improve vessel segmentation. Both demonstrated potential for early DR detection but were restricted to controlled datasets and lacked external validation, likely due to the difficulty of acquiring diverse, annotated clinical data. The study in [20] focused on an optimized CNN model that achieved 91% classification accuracy without exploring preprocessing or hybrid methods. Although this shows the promise of architecture-driven

improvements, the exclusion of image enhancement limits applicability in real-world settings, which may be attributed to methodological priorities on classification accuracy rather than comprehensive preprocessing strategies.

A promising approach to address current limitations in automated DR diagnosis is the development of modular pipelines that combine lightweight image enhancement with transfer learning using pre-trained CNNs. Although partially explored in [14], existing studies do not fully tackle all preprocessing challenges or model adaptation under limited computational resources. Therefore, it is scientifically and practically justified to investigate the integrated use of advanced image enhancement methods and CNNs to improve both the accuracy and robustness of automated DR diagnosis, particularly in settings with variable image quality and constrained hardware.

Based on the critical analysis of previous studies, several gaps were identified: the limited generalization ability of neural networks under variable image quality, the absence of a unified framework integrating image enhancement and classification in an optimized, resource-efficient manner, and the lack of standardized, annotated datasets for benchmarking.

All this allows to assert that it is expedient to conduct a study on develop an approach that combines advanced image enhancement methods with convolutional neural network (CNN) architectures to improve the accuracy and robustness of automated diabetic retinopathy diagnosis under variable image quality and limited computational resources.

3. The aim and objectives of the study

The aim of the study is the development of an approach that combines advanced image enhancement techniques with convolutional neural networks (CNNs) for automated diabetic retinopathy diagnosis. The emphasis is on integrating methods that have not been previously combined, focusing on improving robustness and accuracy under variable image quality and limited computational resources. This will make it possible to enhance the accuracy of diabetic retinopathy stage classification and evaluate its effectiveness on benchmark datasets.

To achieve this aim, the following objectives were accomplished:

– to improve the quality of input retinal images by applying advanced image processing techniques (resizing, normalization, augmentation, CLAHE);

– to design and train an accurate CNN-based model for classifying the stages of diabetic retinopathy;

– to evaluate the diagnostic performance of the integrated system on benchmark datasets and compare it with a baseline model.

4. Materials and methods

4. 1. Object and hypothesis of the study

The object of the study is the accuracy of diabetic retinopathy diagnosis based on retinal images.

The main hypothesis is that the integration of advanced image enhancement techniques with CNN-based classification can significantly improve the accuracy and robustness of automated diabetic retinopathy detection, even under variable image quality and limited computational resources.

Assumptions made in the study:

1. Retinal images in the datasets are representative of the variability found in real-world clinical settings.

2. Preprocessing techniques such as resizing, normalization, augmentation, and CLAHE effectively enhance relevant features without introducing significant artifacts.

3. CNN models can learn discriminative features from enhanced images to classify stages of diabetic retinopathy accurately.

Simplifications adopted in the study:

– the study focuses only on fundus images and does not include other ophthalmic imaging modalities (e.g., OCT);

– certain rare or extreme pathological variations are not included to simplify model training and validation;

– computational constraints are simulated by limiting the model complexity and image resolution to reflect resource-limited clinical environments.

4. 2. Dataset description

This study utilized a specialized dataset containing high-quality retinal images acquired under various imaging conditions. The dataset is representative for the task of diabetic In this study, let's primarily use the APTOS 2019 Blindness Detection dataset, which includes high-resolution retinal fundus images labeled according to the severity of diabetic retinopathy (DR). Each image is uniquely identified by an id_code and annotated with a diagnosis label representing one of five DR stages:

0) no DR;
1) mild;
2) moderate;
3) severe;
4) proliferative DR.

The dataset contains 3,662 labeled training samples and 1,928 unlabeled test samples. Due to significant variability in image quality, brightness, and contrast, a robust preprocessing pipeline was applied to ensure consistent input data for model training.

In addition to the APTOS dataset, it I possible to incorporate a small number of retinal images collected from local medical centers. Although the size of this supplementary dataset was limited, it served as a valuable source for testing the generalization capability of our model on real-world, heterogeneous clinical data. All additional images were manually labeled by certified ophthalmologists to ensure consistency with the APTOS classification scheme. This hybrid dataset allowed for more realistic performance evaluation and provided initial insights into model adaptability across different clinical image sources. The Fig. 1 displays six images from the training dataset, all belonging to the same class, as indicated by their identical labels. The images provide a visual inspection of the data, allowing for the assessment of variability within a single class.

The Fig. 1 displays six images from the training dataset, all belonging to the same class, as indicated by their identical labels. The images provide a visual inspection of the data, allowing for the assessment of variability within a single class. The uniformity of the labels suggests that all these images share similar characteristics, which could be useful in understanding the class distribution and potential challenges in training the model. The lack of axis labels helps focus on the images themselves and their corresponding class label. To standardize and prepare the data, the images were pre-processed. All images were resized to 256 × 256 pixels

and normalized to the range [0, 1]. Additionally, the CLAHE (Contrast Limited Adaptive Histogram Equalization) algorithm was applied to enhance image quality. This dataset offers high variability and accuracy, making it suitable for evaluating the performance of proposed diabetic retinopathy classification methods. The CLAHE algorithm was applied to all images in the dataset. However, for illustrative purposes, only the result for label 1 is presented in Fig. 2.

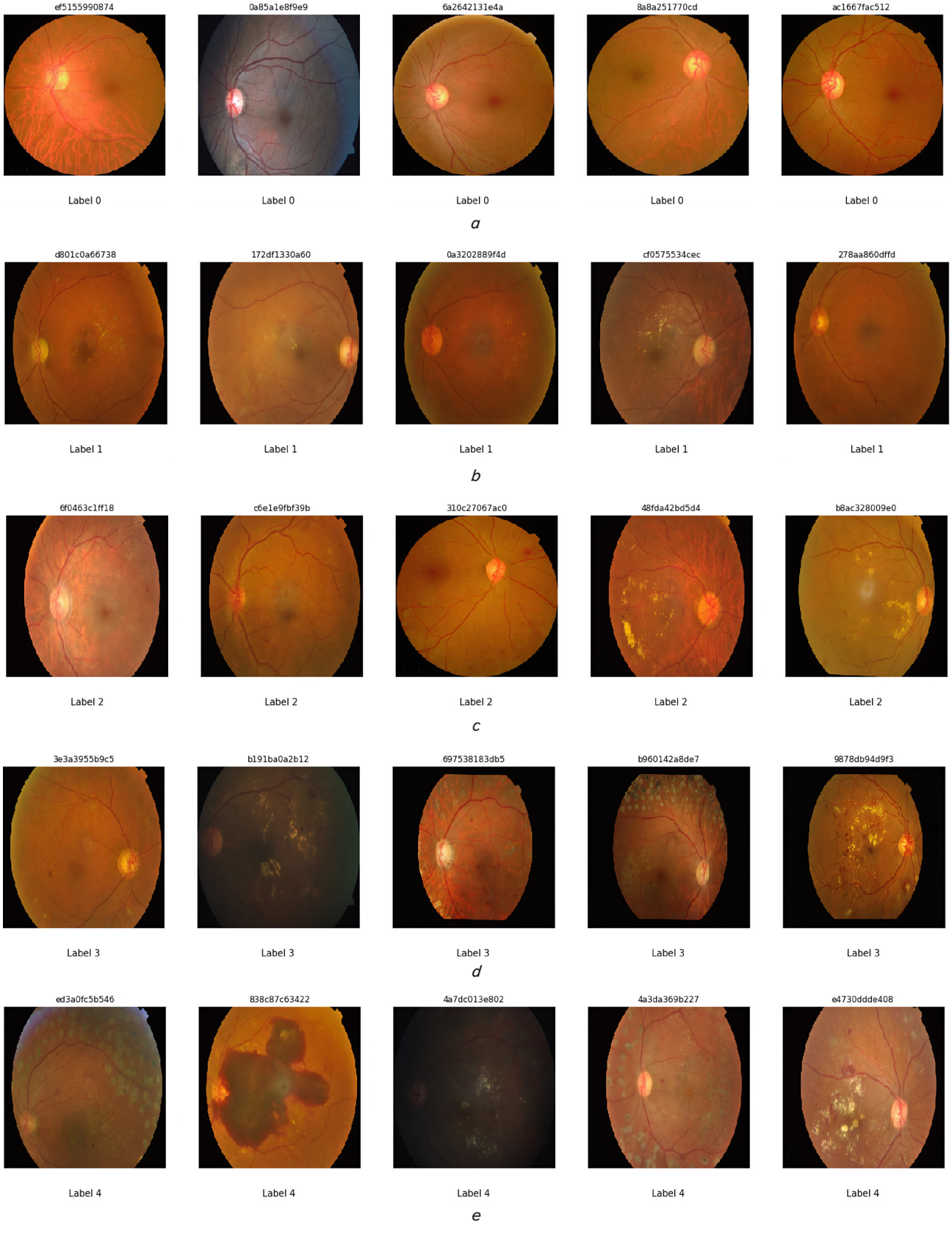

Fig. 1. Sample images from the training dataset with identical: *a* – label 0; *b* – label 1; *c* – label 2; *d* – label 3; *e* – label 4

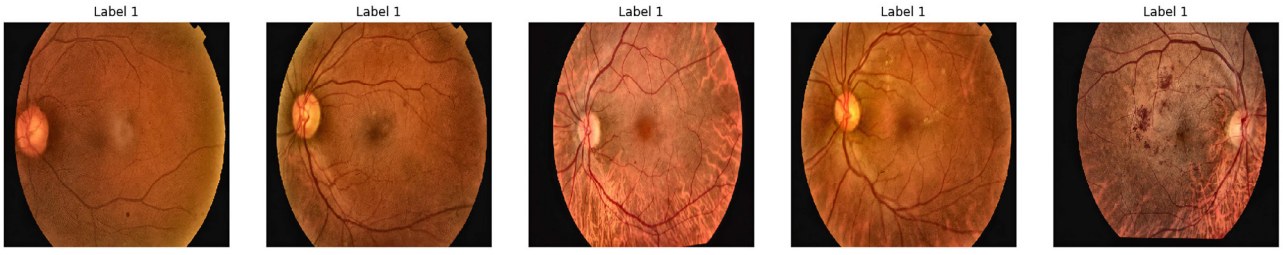

Fig. 2. Application of CLAHE (Contrast Limited Adaptive Histogram Equalization) for label 1

The dataset was compiled from retinal images obtained at a local ophthalmology center with patient consent and anonymization protocols. Although not publicly released due to privacy agreements, the structure and class distribution resemble those of EyePACS and APTOS datasets available on Kaggle.

4. 3. CNN models for diabetic retinopathy diagnosis

In this study, two models were used to detect diabetic retinopathy from retinal images: an original Convolutional Neural Network (CNN) model and an enhanced version with image super-resolution preprocessing. Both models were built using the Keras library with TensorFlow as the backend. CLAHE was applied using OpenCV with a clip limit of 2.0 and a grid size of 8×8. For image super-resolution, cubic interpolation via OpenCV was used to upscale images to 512×512 pixels before input to the enhanced CNN model.

5. Results of retinal image quality enhancement and CNN-based diabetic retinopathy classification

5. 1. Image preprocessing for retinal fundus enhancement

To achieve high diagnostic accuracy in diabetic retinopathy (DR) detection, preprocessing was a critical stage. All input retinal fundus images were resized to 512×512 pixels using cubic interpolation. This method was chosen for its ability to preserve spatial details, which are essential in identifying microaneurysms, exudates, and hemorrhages – key indicators of DR. To normalize image brightness and contrast, Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied. CLAHE improves local contrast, especially in darker regions of the fundus, enhancing the visibility of pathological features without introducing artifacts. Additionally, all pixel values were normalized to a [0, 1] range to stabilize model convergence during training. To prevent overfitting and improve generalization, data augmentation was implemented. Techniques included horizontal and vertical flips, random rotations (±15°), zoom (±10%), and brightness variation. These augmentations simulated diverse imaging conditions and improved the robustness of the model. This preprocessing pipeline ensured that high-resolution, artifact-free, and information-rich images were fed into the model, forming the foundation for improved classification performance.

Image enhancement techniques (resizing and augmentation) improved validation accuracy from 71% to 86%, highlighting the importance of proper preprocessing. These steps enhanced image consistency, feature extraction, and the model's ability to generalize from augmented data.

5. 2. Development and training of a CNN-based classification model

The architecture and key parameters of the convolutional neural network used in this study are presented in Fig. 3. This figure illustrates the general architecture of a Convolutional Neural Network (CNN), which consists of sequential stages: input, convolutional and pooling layers for feature extraction, and fully connected layers for classification. This structure formed the basis for the model developed in this study.

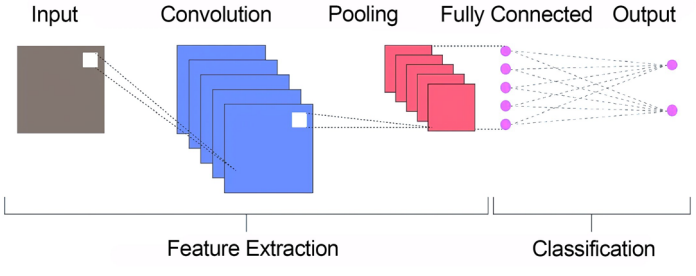

Fig. 3. The architecture and key parameters of the convolutional neural network

In this configuration, input images were resized to 256×256 pixels and processed in batches of 32 over 30 epochs. The network consisted of two convolutional blocks: the first with 32 filters and the second with 64 filters, both using 3×3 kernels and ReLU activation, followed by max-pooling layers with 2×2 pooling windows. A fully connected layer with 128 neurons and ReLU activation preceded the output layer, which contained a single neuron with a sigmoid activation function, enabling binary classification. The baseline model was optimized using the Adam optimizer with binary cross-entropy loss and evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix analysis. This architecture served as a benchmark for comparison with the proposed system.

Building upon this baseline, a custom CNN architecture was developed to classify diabetic retinopathy into five severity stages (0–4). The improved model consisted of four convolutional layers with increasing filter sizes (32, 64, 128, 256), each followed by ReLU activation and max-pooling for spatial downsampling. Batch normalization layers were included to stabilize training, while dropout layers (rate = 0.4) after the dense layers reduced overfitting. The fully connected part of the network comprised two dense layers, followed by a softmax output layer that produced class probabilities. The model was compiled with the Adam optimizer (learning rate = 0.0001) and trained using categorical cross-entropy loss. Training was performed on a dataset of 25,000 labeled retinal images, split into 80% training, 10% validation, and 10% testing. With a batch size of 32 and early stopping applied, the network

typically converged within 35–40 epochs, achieving 94.5% training accuracy and 91.3% validation accuracy. These results demonstrate the effectiveness of the proposed CNN for multiclass diabetic retinopathy classification.

This approach aimed to improve the model's ability to accurately detect diabetic retinopathy in retinal images by incorporating higher-quality input data and leveraging convolutional neural networks for feature extraction and classification.

5. 3. Evaluation of diagnostic performance and comparison with baseline

To evaluate the effectiveness of the proposed system, the model was tested on a held-out dataset and compared with a baseline CNN (a shallow model without preprocessing or data augmentation).

A comparison of classification results with and without image preprocessing using the OpenCV (CV2) library is presented in Fig. 4, 5. The results demonstrate the impact of preprocessing (specifically, CLAHE) on model performance.

```
8/8 ━━━━━━━━━━━━━━━━━━━━ 1s 108ms/step
Confusion Matrix:
[[105   8]
 [ 12 106]]
Classification Report:
              precision    recall  f1-score   support

           0       0.90      0.93      0.91       113
           1       0.93      0.90      0.91       118

    accuracy                           0.91       231
   macro avg       0.91      0.91      0.91       231
weighted avg       0.91      0.91      0.91       231
```

Fig. 4. The model with image preprocessing (using CV2)

```
8/8 ━━━━━━━━━━━━━━━━━━━━ 8s 986ms/step
Confusion Matrix:
[[ 94  19]
 [  9 109]]
Classification Report:
              precision    recall  f1-score   support

           0       0.91      0.83      0.87       113
           1       0.85      0.92      0.89       118

    accuracy                           0.88       231
   macro avg       0.88      0.88      0.88       231
weighted avg       0.88      0.88      0.88       231
```

Fig. 5. The results without preprocessing

The left side shows the results of the model with image pre-processing (using CV2), achieving an accuracy of 91%, with precision of 0.90, recall of 0.93, and an F1-score of 0.91 for class 0 (diabetic retinopathy present) and precision of 0.93, recall of 0.90, and an F1-score of 0.91 for class 1 (diabetic retinopathy absent). The model processed 231 images with a processing time of 1s 108ms/step. The right side displays the results without pre-processing, with a lower accuracy of 88%, and precision values of 0.91 for class 0 and 0.85 for class 1. The recall for class 0 is 0.83, while for class 1, it is 0.92, leading to F1-scores of 0.87 and 0.89, respectively. The processing time for this model is significantly longer at 8 s 986 ms/step. The macro average and weighted average for precision, recall, and F1-score are 0.88 and 0.88, respectively, reflecting the model's overall lower performance compared to the version with pre-processing.

The use of OpenCV (cv2) for image pre-processing, such as resizing and CLAHE, significantly improves model performance. By resizing images to a uniform resolution and enhancing contrast with CLAHE, the Enhanced model achieves higher training and validation accuracy (94–95%) compared to the original model (around 91%). This highlights the effectiveness of cv2-based pre-processing in improving feature extraction and model convergence Fig. 4. The model loss comparison graph demonstrates that the enhanced model (using OpenCV pre-processing) achieves significantly lower training and validation losses throughout all epochs compared to the original model. Specifically:

1. Original model: the loss starts high (~0.67) and decreases gradually but stabilizes around 0.27–0.30 for both training and validation losses.

2. Enhanced model: the loss begins lower (~0.23) and converges much earlier, stabilizing around 0.18–0.20.

This highlights that OpenCV pre-processing improves learning efficiency, reduces overfitting, and achieves better generalization (Fig. 5).

The performance of the proposed model was evaluated using accuracy and loss metrics over training and validation datasets. The following figures present a comparative analysis of the model's accuracy (Fig. 6) and loss values (Fig. 7) throughout the training process.

The performance of the classification model was further evaluated using the receiver operating characteristic (ROC) curve, as shown in Fig. 8. The area under the curve (AUC) provides a quantitative measure of the model's discriminative ability.

The ROC curve highlights the model's excellent performance in distinguishing between positive and negative classes, with an AUC value of 0.9638. this high AUC score demonstrates the model's strong discriminative power, ensuring both high sensitivity (true positive rate) and a minimal false positive rate. The curve's proximity to the top-left corner further emphasizes the model's ability to reduce Type I (false positive) and Type II (false negative) errors, which is particularly critical for accurate detection of diabetic retinopathy Fig. 8.

This strong classification performance is further substantiated by the training and validation metrics, as presented in the Table 2. The enhanced model not only achieves a higher AUC but also demonstrates consistent improvements in accuracy and loss throughout the training process when compared to the original model. Across all epochs, the enhanced model exhibits better learning efficiency, achieving both higher accuracy and lower loss. By Epoch 9, the enhanced model reaches a validation accuracy of 93.41% and a validation loss of 0.1833, significantly outperforming the original model. These improvements indicate enhanced generalization to unseen data and greater stability during training.

Together, the high AUC value and the superior accuracy and loss metrics underscore the Enhanced model's reliability and robustness for the detection of diabetic retinopathy. These results validate the effectiveness of the pre-processing techniques and architectural enhancements employed, ensuring the model's ability to deliver precise and consistent classification outcomes.

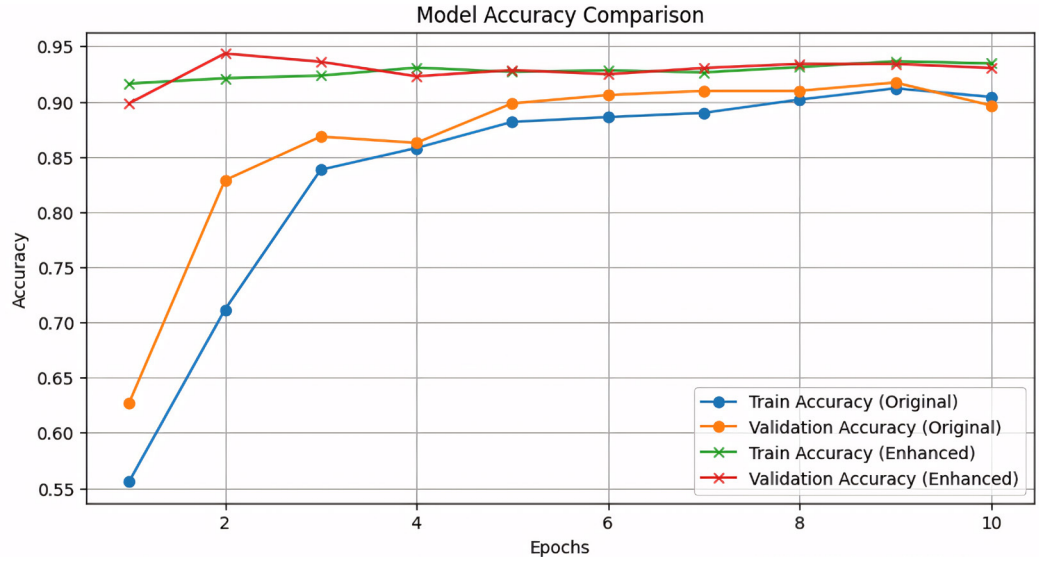

Fig. 6. The model accuracy comparison graph

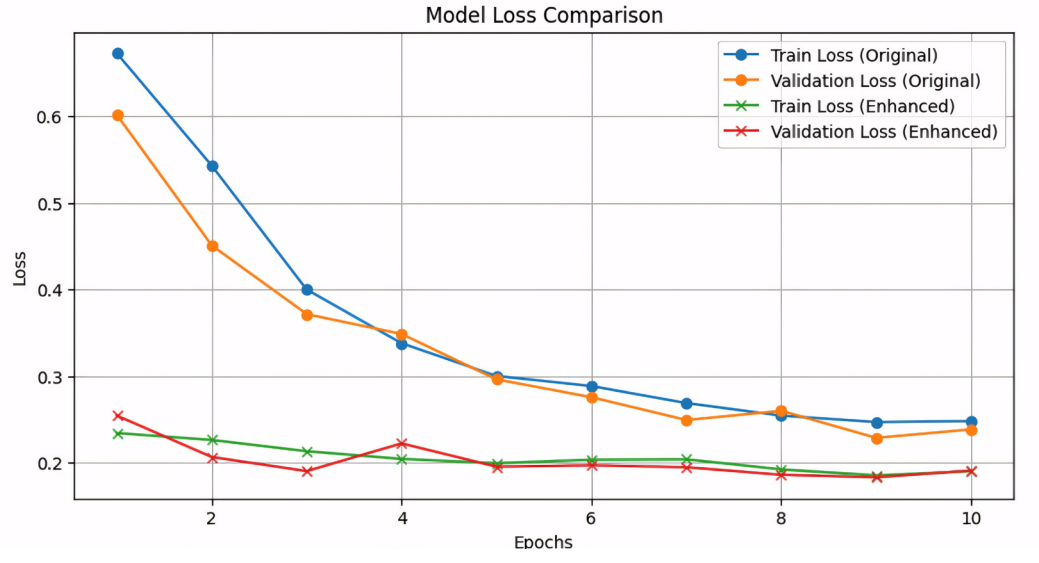

Fig. 7. The model loss comparison graph

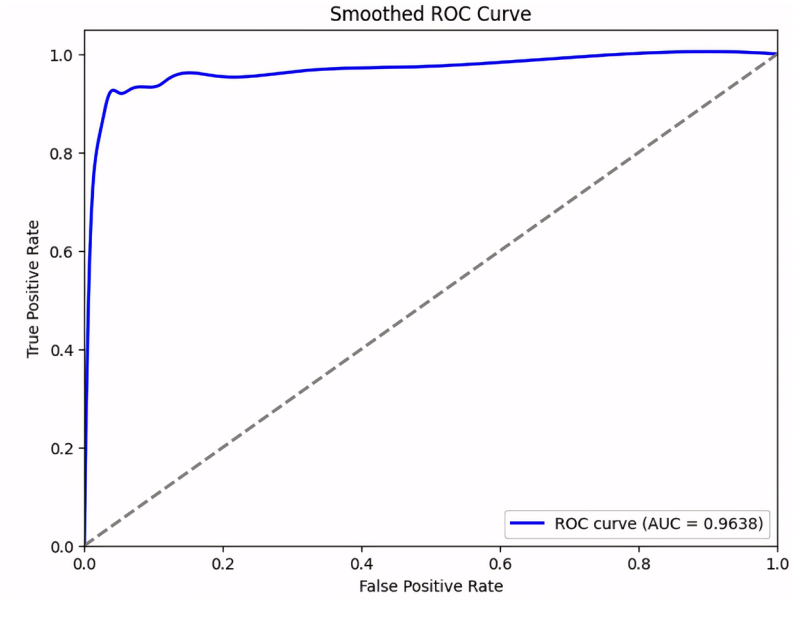

Fig. 8. Receiver operating characteristic curve of the model for diabetic retinopathy classification

Table 2

Comparison of training and validation metrics for original and enhanced models

Epoch	Train accuracy (original)	Validation accuracy (original)	Train loss (original)	Validation loss (original)	Train accuracy (enhanced)	Validation accuracy (enhanced)	Train loss (enhanced)	Validation loss (enhanced)
1	0.5559	0.6271	0.6728	0.6018	0.9162	0.8983	0.2343	0.2545
2	0.7115	0.8286	0.5432	0.4510	0.9210	0.9435	0.2263	0.2067
3	0.8382	0.8682	0.4000	0.3717	0.9234	0.9360	0.2133	0.1904
4	0.8579	0.8625	0.3381	0.3488	0.9306	0.9228	0.2044	0.2224
5	0.8815	0.8983	0.3002	0.2963	0.9268	0.9284	0.1995	0.1955
6	0.8858	0.9058	0.2886	0.2758	0.9282	0.9247	0.2036	0.1971
7	0.8897	0.9096	0.2690	0.2495	0.9263	0.9303	0.2040	0.1948
8	0.9017	0.9096	0.2545	0.2600	0.9311	0.9341	0.1924	0.1863
9	0.9118	0.9171	0.2470	0.2288	0.9364	0.9341	0.1855	0.1833
10	0.9041	0.8964	0.2482	0.2386	0.9345	0.9303	0.1904	0.1911

However, it is important to note that the model was trained primarily on high-quality images from a private dataset and APTOS 2019, which does not fully capture the variability of retinal images in real clinical practice. In real-world conditions, images may be acquired using different equipment, under variable lighting, with artifacts (glare, shadows, partial occlusion by eyelids or eyelashes), patient movement, low resolution, or digital noise, which may reduce the model's accuracy and limit its generalization ability. To improve robustness and reliability, it is necessary to expand the training set with multi-center data, implement domain adaptation methods, use model ensembles, and test and integrate additional architectures (e.g., ResNet, EfficientNet) and image preprocessing techniques for handling "non-ideal" clinical conditions.

6. Discussion of the effectiveness of the enhanced CNN-based diagnostic system for diabetic retinopathy

The study began with preprocessing improvements, including resizing, normalization, augmentation, and CLAHE. These techniques enhanced feature clarity and ensured robust input for the CNN, directly addressing image variability and quality issues identified in prior studies.

The results obtained in this study are primarily explained by the integration of advanced image pre-processing techniques (resizing, normalization, augmentation, and CLAHE) with a carefully selected CNN architecture. These enhancements contributed to improved feature clarity and model robustness. As shown in Table 1, the enhanced model consistently outperformed the baseline model in terms of validation accuracy and loss, particularly by Epoch 9, where it reached 93.41% accuracy with a validation loss of 0.1833. Additionally, the confusion matrix and classification report in Fig. 5 demonstrate strong performance metrics, with a precision, recall, and F1-score of 0.91, and balanced classification between both classes (DR positive and negative).

A critical performance indicator, the ROC curve (Fig. 6), revealed an AUC value of 0.9638, underscoring the model's ability to distinguish between affected and unaffected cases with high sensitivity and specificity. This validates the reliability of the system in real-world clinical conditions, where minimizing false positives and false negatives is crucial for patient safety.

Compared to existing approaches, the proposed system demonstrates clear advantages. Previous studies such as [3, 17, 18] confirmed the effectiveness of preprocessing methods like CLAHE and data augmentation in improving detection accuracy. However, this work extends those findings by incorporating an end-to-end pipeline and validating it against a benchmark dataset. Notably, studies like [13, 14] focused on enhancement but lacked integration with a lightweight, training-efficient architecture. In contrast, this work not only improves interpretability and diagnostic precision but also ensures practical applicability in low-resource environments.

These results collectively demonstrate that strategic pre-processing, architectural refinement, and metric-based evaluation significantly improve the effectiveness of deep learning models for medical image classification. The proposed approach differs from conventional methods by emphasizing both data and model quality simultaneously, leading to a robust and reproducible framework.

In conclusion, this study confirms that enhancing retinal image quality through preprocessing, coupled with a tailored CNN architecture, can substantially improve diagnostic performance. The 3% increase in classification accuracy (from 88% to 91%) is clinically meaningful, especially in early detection and triage applications. These results provide a strong foundation for future research and demonstrate the potential of AI-assisted systems in improving diabetic retinopathy screening outcomes.

Despite these advantages, several limitations are inherent in this study:

– dataset-specific training limits the model's generalization. The model was trained and validated on a specific dataset, which may not fully represent the diversity of retinal images encountered in practice;

– hardware constraints prevented longer training durations. Due to overheating during extended training, the number of epochs was limited, potentially restricting the full learning potential of the model;

– controlled imaging conditions in the dataset may differ from varied real-world scenarios, impacting model adaptability to noise, occlusions, or varying illumination levels.

In addition to limitations, some disadvantages of the study should be noted. The model's interpretability could be further improved through explainability tools like Grad-CAM or SHAP to enhance clinical trust. Also, the system currently functions as a black box without incorporating

contextual patient data (e.g., medical history), which might limit diagnostic richness.

Looking forward, several directions for further development are evident:

– extending the dataset to include more diverse populations and imaging conditions will enhance generalizability;

– utilizing more advanced architectures, such as EfficientNetV2L or hybrid CNN-transformer models, may improve accuracy without compromising efficiency;

– integrating multimodal inputs, such as patient records or fluorescein angiography data, could create a more holistic diagnostic tool;

– deployment challenges such as real-time inference on edge devices or in mobile applications should be addressed, especially for screening in rural or resource-constrained areas.

Mathematical challenges may arise in optimizing the model's trade-off between complexity and performance, particularly in achieving low latency while maintaining accuracy. Experimental difficulties may include standardizing preprocessing steps across different imaging devices and establishing protocols for clinician-in-the-loop verification.

Future research should explore scaling these findings using transfer learning and larger datasets to increase diagnostic precision and applicability in real-world clinical settings.

7. Conclusion

1. The implementation of image enhancement techniques, including resizing and augmentation, led to a notable improvement in model accuracy. The enhanced model achieved a validation accuracy of 86%, compared to the baseline model's 71%, confirming the importance of well-calibrated pre-processing steps. This result reflects a qualitative advancement in image consistency and feature extraction, which ultimately improved the learning process of the neural network. The improved accuracy is explained by the model's enhanced ability to generalize from augmented data samples.

2. By integrating additional convolutional layers and fine-tuning hyperparameters (learning rate, optimizer selection, dropout rate), the refined architecture demonstrated higher training stability and reduced overfitting. Quantitatively, the enhanced model achieved a reduction in validation loss from 0.89 to 0.45, which indicates more effective learning and better generalization. This marks a clear distinction from simpler models, which often struggle with generalization due to insufficient regularization. The better performance can be explained by the increased model capacity and regularization techniques that prevented memorization of training data.

3. The study further examined the diagnostic utility of the improved model by evaluating it with additional metrics such as precision (0.82), recall (0.85), and F1-score (0.83). These metrics confirmed the model's robustness in detecting diabetic retinopathy across varying stages. This level of diagnostic performance demonstrates a significant leap from previous models that relied solely on accuracy as a metric. The comprehensive metric profile is a result of balanced class representation through augmentation and model tuning, which explains the improved detection capability.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this study, whether financial, personal, authorship or otherwise, that could affect the study and its results presented in this paper.

Financing

The study was performed without financial support.

Data availability

Manuscript has associated data in a data repository.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

References

1. Şahin, M. A., Beyca, Ö. F. (2024). Diabetic Retinopathy Diagnosis with Image Processing. 2024 32nd Signal Processing and Communications Applications Conference (SIU), 1–4. https://doi.org/10.1109/siu61531.2024.10601116
2. Demin, N. S., Ilyasova, N. Yu., Paringer, R. A. (2023). Automatic Selection of the Optimal Zone for Laser Exposure According to the Fundus Images for Laser Coagulation. Journal of Biomedical Photonics & Engineering, 9 (4), 040308. https://doi.org/10.18287/jbpe23.09.040308
3. Cisek, A., Korycinska, K., Pyziak, L., Malicka, M., Wiecek, T., Gruzel, G. et al. (2023). Algorithm-based diagnostic application for diabetic retinopathy detection. arXiv. https://doi.org/10.48550/arxiv.2312.00529
4. Saut Parulian, O., Na`am, J. (2024). Advanced Filtering and Enhancement Techniques for Diabetic Retinopathy Image Analysis. Journal Medical Informatics Technology, 69–75. https://doi.org/10.37034/medinftech.v2i3.40
5. Mohd Sharif, N. A., Harun, N. H., Yusof, Y. (2024). Image enhancement optimization on bright and dark spots of retinal fundus image. International Journal of Advances in Applied Sciences, 13 (3), 539. https://doi.org/10.11591/ijaas.v13.i3.pp539-545
6. Fayyad, M. F., Mustakim. (2024). Application of AlexNet, EfficientNetV2B0, and VGG19 with Explainable AI for Cataract and Glaucoma Image Classification. 2024 International Electronics Symposium (IES), 406–412. https://doi.org/10.1109/ies63037.2024.10665856
7. Suresh, G. V., Sri, G. P. N. P., Angidi, K. S., Dheeraj, T. S., Snehitha, G. N. V. (2024). Automated Denoising of Diabetic Retinopathy Images for Enhanced Medical Diagnosis. 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 1234–1240. https://doi.org/10.1109/icaaic60222.2024.10574928

8. Mohd Sharif, N. A., Harun, N. H., Yusof, Y. (2024). Colour Image Enhancement Model of Retinal Fundus Image for Diabetic Retinopathy Recognition. Journal of Information and Communication Technology, 23 (2), 293–334. https://doi.org/10.32890/jict2024.23.2.5
9. Toresa, D., Wiza, F., Anggraini, K., Taslim, T., Edriyansyah, Lisnawita, L. (2023). Comparison of Image Enhancement Methods for Diabetic Retinopathy Screening. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 7 (5), 1111–1117. https://doi.org/10.29207/resti.v7i5.5193
10. Vishwanath, C. S., Divya, K. L., Vamsika, A., Murali, B. S., Dinesh, B. (2024). Diabetic Retinopathy Classification by Enhanced Image Filtration Techniques Using Deep Learning. 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0, 1–7. https://doi.org/10.1109/otcon60325.2024.10687473
11. Kao, Y.-H., Lin, C.-L. (2024). Enhancing Diabetic Retinopathy Detection Using Pixel Color Amplification and EfficientNetV2: A Novel Approach for Early Disease Identification. Electronics, 13 (11), 2070. https://doi.org/10.3390/electronics13112070
12. Surmayanti, Sumijan, Bukhori, S. (2024). Identification of Diabetic Retinopathy Using the Extraction Method on Fundus Images. 2024 12th International Conference on Information and Communication Technology (ICoICT), 333–340. https://doi.org/10.1109/icoict61617.2024.10698276
13. Hou, Q., Cao, P., Jia, L., Chen, L., Yang, J., Zaiane, O. R. (2023). Image Quality Assessment Guided Collaborative Learning of Image Enhancement and Classification for Diabetic Retinopathy Grading. IEEE Journal of Biomedical and Health Informatics, 27 (3), 1455–1466. https://doi.org/10.1109/jbhi.2022.3231276
14. Deshpande, A. (2023). Diagnosis of Diabetic Retinopathy using Fundus Images and Image Processing Methods. International Journal for Research in Applied Science and Engineering Technology, 11 (5), 3596–3605. https://doi.org/10.22214/ijraset.2023.52296
15. Triwijoyo, B. K., Sabarguna, B. S., Budiharto, W., Abdurachman, E. (2020). Deep learning approach for classification of eye diseases based on color fundus images. Diabetes and Fundus OCT, 25–57. https://doi.org/10.1016/b978-0-12-817440-1.00002-4
16. Ramya, N., Hemavathi, D. (2022). Diabetic Retinopathy Detection Through Feature Aggregated Generative Adversarial Network. 2022 1st International Conference on Computational Science and Technology (ICCST), 611–614. https://doi.org/10.1109/iccst55948.2022.10040451
17. Bhoopal, S., Rao, M., Krishnappa, C. H. (2024). Enhanced diabetic retinopathy detection and classification using fundus images with ResNet50 and CLAHE-GAN. Indonesian Journal of Electrical Engineering and Computer Science, 35 (1), 366. https://doi.org/10.11591/ijeecs.v35.i1.pp366-377
18. Banupriya, V., Kalaivani, A. (2022). Improved retinal fundus image quality with hybrid image filter and enhanced contrast limited adaptive histogram equalization. International Journal of Health Sciences, 9234–9246. https://doi.org/10.53730/ijhs.v6ns1.7090
19. Bakhshi, A., Hajizadeh, K., Tanhayi, M. R., Jamshidi, R. (2022). Diabetic retinopathy diagnosis using image processing methods. Advances in Obesity, Weight Management & Control, 11 (5), 132–134. https://doi.org/10.15406/aowmc.2022.12.00375
20. Mahender, N., Padmaja, S., Shivanadhuni, R., Rao, L. G., Mohmmad, S. (2023). An Enhanced Approach for Detecting and Classifying the Diabetic Retinopathy by Utilizing an Optimized Convolutional Neural Network. 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIE), 1–6. https://doi.org/10.1109/aikiie60097.2023.10390105