A Customized CNN Architecture with CLAHE for Multi-Stage Diabetic Retinopathy Classification

Songgrod Phimphisan
Department of Computer Science and Information Technology, Faculty of Science and Health Technology, Kalasin University, Thailand
songgrod.ph@ksu.ac.th

Nattavut Sriwiboon
Department of Computer Science and Information Technology, Faculty of Science and Health Technology, Kalasin University, Thailand
nattavut.sr@ksu.ac.th (corresponding author)

Received: 6 September 2024 | Revised: 1 October 2024 | Accepted: 16 October 2024

 DOI: https://doi.org/10.48084/etasr.8932

ABSTRACT

This paper presents a customized Convolutional Neural Network (CNN) architecture for multi-stage detection of Diabetic Retinopathy (DR), a leading cause of vision impairment and blindness. The proposed model incorporates advanced image enhancement techniques, particularly Contrast Limited Adaptive Histogram Equalization (CLAHE), to improve the visibility of critical retinal features associated with DR. By integrating CLAHE with a finely tuned CNN, the proposed approach significantly enhances accuracy and robustness, allowing for more precise detection across various stages of DR. The proposed model was evaluated against several state-of-the-art techniques, with the customized CNN alone achieving an overall accuracy of 97.69%. The addition of CLAHE further boosts the performance, achieving an accuracy of 99.69%, underscoring the effectiveness of combining CLAHE with CNN for automated DR detection. This approach provides an efficient, scalable, and highly accurate solution for early and multistage DR detection, which is crucial for timely intervention and prevention of vision loss.

Keywords-diabetic retinopathy; multiple stages; CNN; CLAHE; classification

I. INTRODUCTION

Diabetic Retinopathy (DR) is a significant risk to public health, leading to blindness and loss of vision [1]. Various studies have been conducted to address early-stage detection and severity classification of DR using deep learning models. Early detection and accurate classification of DR stages are crucial for preventing vision impairment and guiding appropriate treatment strategies. Fundus imaging is a widely used non-invasive technique for diagnosing and monitoring the progression of DR. Manual analysis of fundus images by clinicians is time-consuming and subject to variability, which can lead to inconsistent diagnoses. Consequently, there is a growing interest in using deep learning techniques, particularly Convolutional Neural Networks (CNNs) [2], to automate the detection and classification of DR stages from fundus images. CNNs have demonstrated exceptional performance in various medical imaging tasks, including the classification of retinal diseases. Numerous studies have proposed the use of deep learning and CNNs for DR classification [3-11]. However, many previous works have encountered challenges, particularly in terms of scalability and efficiency, indicating a need for further optimization. To address these challenges, Contrast Limited Adaptive Histogram Equalization (CLAHE) [12] has emerged as a powerful preprocessing technique to enhance image contrast, making subtle features more discernible. This paper proposes a combined approach that integrates a customized CNN architecture with a customized CLAHE preprocessing technique for the classification of DR stages. Enhancing the contrast of fundus images using a customized CLAHE, followed by feature extraction and classification using a specialized CNN, aims to improve the accuracy and reliability of DR stage detection. The contributions of this study are as follows.

- Combines a custom CNN architecture and CLAHE to classify multi-stage DR.
- Presents a high-performance approach for DR classification.
- Achieves higher accuracy compared to seven other previous works.

II. BACKGROUND AND RELATED WORKS

A. Convolutional Neural Network (CNN)

CNNs are highly effective in medical image analysis. CNNs automatically learn features from complex medical images. These networks consist of layers such as convolutional, pooling, and fully connected layers, which progressively extract and analyze visual patterns in the images. Each time a convolution is performed, a new convolved image is generated, capturing features extracted from the previous layer's image. If $I(x,y)$ represents a 2D input image and $f(x,y)$ denotes the 2D kernel used for the convolution, the process can be described as follows [13]:

$$y(x,y) = (I,f)(x,y) = \sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty} I(x-u,y-v)f(u,v) \quad (1)$$

During the convolution process, edge pixel values can either be disregarded, or padding can be added to include them. The result of the convolution can then be modified using a nonlinear activation function [14].

$$sigmoid(x) = \frac{1}{1+e^{-x}} \quad (2)$$

In addition to convolutional layers, CNNs have pooling and fully connected layers. Pooling layers reduce the size of feature maps by downsampling, using max or average pooling to condense local regions. This process includes pixel spacing, called stride. Unlike other layers, pooling layers don't have activation functions but often use ReLU for non-linearity. The average pooling is calculated for each convolutional layer as needed [15].

$$X_{ij}^{[l]} = \frac{1}{MN}\sum_{m}^{M}\sum_{n}^{N} X_{iM+m,jN+n}^{[l-1]} \quad (3)$$

In this context, i and j represent the coordinates of the output map and M and N denote the sizes of the pooling samples. To convolutional and pooling layers, CNNs use fully connected layers for classification. The extracted features are passed to these layers, where each connection has its own weight, requiring significant computational power. The standard sigmoid function is applied for final classification [15]:

$$S(t) = \frac{1}{1+e^{-x}} \quad (4)$$

B. Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE enhances local image contrast by applying histogram equalization to small blocks and clipping the histogram to control contrast and reduce noise [16-17]. CLAHE improves feature visibility without introducing artifacts, making it ideal for detailed image enhancement. The mathematical formulation of CLAHE is as follows: let $I(x,y)$ represent the intensity of the pixel at location (x,y) in the image. The image is divided into small non-overlapping blocks or tiles of size $M \times N$. For each tile, the histogram $H(i)$ is computed, where i represents the intensity level. A clip limit T is defined. If any histogram bin $H(i)$exceeds T, it is clipped:

$$H_{clipped}(i) = min(H(i),T) \quad (5)$$

The excess pixels above the clip limit T are then redistributed across all histogram bins. Normalize the Cumulative Distribution Function (CDF) to obtain the intensity mapping function. The CDF for the clipped histogram is calculated as:

$$CDF(i) = \sum_{j=0}^{i} H_{clipped}(j) \quad (6)$$

To avoid artifacts between neighboring tiles, the intensity value for a pixel is interpolated using the CDFs of the surrounding tiles. The final pixel value after applying CLAHE is given by:

$$I_{CLAHE}(x,y) = CDF\big(I(x,y)\big) \times \frac{MaxIntensity - MinIntensity}{TotalPixelsinTile} \quad (7)$$

where the $MaxIntensity$ and $MinIntensity$ are the maximum and minimum intensity values in the image, and $TotalPixelsinTile$ is the number of pixels in each tile.

C. Related Works

Several studies [3-9, 18-20] have applied deep learning techniques to enhance DR detection from retinal fundus images. In [3], the Beluga Whale Optimizer (BWO) was combined with deep learning to assist DR diagnosis through ShuffleNet-v2 for feature extraction and Deep Stacked Autoencoder (DSAE) for classification, achieving 99.58% accuracy. In [4], a hybrid approach was introduced using NASNetLarge for feature extraction and Glowworm Swarm Optimization (GSO) for tuning, employing Variational Autoencoders (VAE) for classification and reaching 99.36% accuracy. In [5], DR grading was improved by integrating bio-inspired optimization with VGGNet and U-Net, resulting in 99.26% accuracy. In [9], transfer learning-based models such as InceptionResNetV2 were optimized for retinal image segmentation, achieving 99.55% accuracy. In [6], the focus was on CNN models enhanced by attention mechanisms and transfer learning, achieving 99.40% accuracy. Heuristic optimizers such as Moth Search were explored in [7], combined with residual learning techniques to refine DR classification and achieve 99.33% accuracy. In [8], lightweight CNNs were developed with ensemble learning, reaching 99.50% accuracy in early DR detection. These studies highlight the importance of hybrid optimization, feature extraction, and classification techniques to improve the accuracy and efficiency of automated DR diagnosis.

III. THE PROPOSED METHOD

CNNs automatically learn features from complex medical images, making them ideal for detecting and classifying abnormalities in fundus images. Figure 1 shows the proposed model approach. This combined approach utilizes a customized CNN architecture and CLAHE to classify the DR stages.

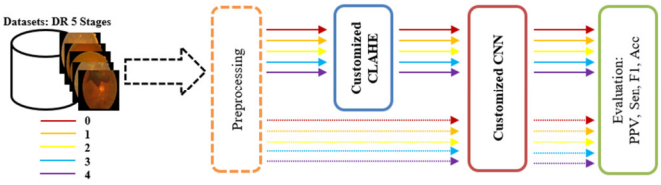

Fig. 1. The proposed approach.

A. Customized CLAHE

The proposed customized CLAHE method enhances the contrast of fundus images through a systematic process that begins with preparing the input image and carefully applying each step based on the mathematical principles of CLAHE. This approach ensures precise and effective contrast enhancement, producing a final output with significantly improved visual clarity. The steps of this process are:

- Prepare the Input Image: Convert the original fundus image to grayscale, simplifying it to focus on pixel intensity values $I(x, y)$. This is essential for CLAHE to effectively enhance brightness and contrast.
- Global Histogram Equalization: Improve the contrast of an image by stretching the intensity range of the pixel values. This redistributes the intensities so that the histogram of the output image is approximately flat.
- Define CLAHE Parameters: Set the clip limit T to 2.0 to control contrast enhancement, and divide the image into non-overlapping tiles of size 8×8 for the tile size $M \times N$. These parameters were chosen to balance detail enhancement with noise suppression.
- Apply Histogram Clipping: Calculate the histogram for each tile and clip any bins exceeding the clip limit $T = 2.0$, redistributing the excess pixels. This step reduces noise while enhancing the contrast within each tile.
- Compute the CDF: Compute the CDF for each tile's clipped histogram, mapping the original pixel values to new intensities. This mapping enhances local contrast, making subtle image details more visible.
- Interpolate Pixel Intensities: Interpolate pixel values using the CDFs from neighboring tiles to ensure smooth transitions between them. This avoids visible seams and creates a cohesive, artifact-free image.
- Generate and Save the Enhanced Image: Generate the final enhanced image with improved contrast and save it. This enhanced image provides better visibility of critical features, aiding in accurate medical analysis.

This customized CLAHE approach modifies the traditional CLAHE method to enhance applications such as DR classification. Figure 2 illustrates the enhancement results of CLAHE.

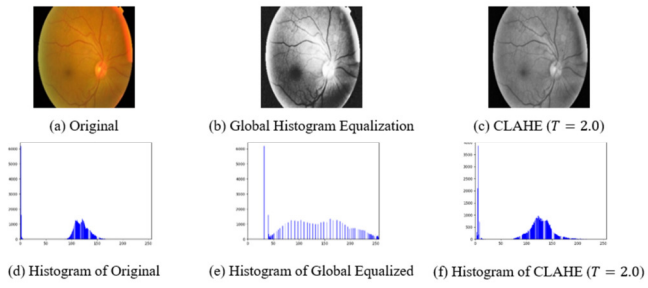

Fig. 2. Enhancement results of CLAHE.

B. Customized CNN Architecture

To simplify the explanation, each set of layers is presented in a section of its own, and the layer-wise parameters are summarized in Table I.

- Input Layer: The input layer receives 224×224 pixel RGB images, feeding them into the network without any alterations. This layer is crucial for establishing the initial data structure but contains no learnable parameters.
- Conv Layer 1: This layer applies 32 filters of size 3×3 with a stride of 1×1, detecting basic features such as edges and textures. The filters create 32 feature maps that capture diverse aspects of the input image.
- Max Pooling 1: This layer reduces the spatial dimensions of the feature maps from Conv Layer 1 using a 2×2 pooling window and a stride of 2×2. It helps decrease computational complexity while maintaining important features.
- Conv Layer 2: This layer uses 64 filters of size 3×3, continuing the feature extraction process by identifying more complex patterns in the input. The stride of 1×1 preserves the spatial dimensions for further processing.
- Max Pooling 2: Similar to Max Pooling 1, this layer downsamples the feature maps from Conv Layer 2, again using a 2×2 window with a stride of 2×2. This step further reduces the size of the feature maps, making the model more efficient.
- Conv Layer 3: With 128 filters of size 3×3, Conv Layer 3 extracts even finer details and more complex features from the image. This layer is key to capturing intricate patterns essential for accurate classification.
- Average Pooling: This layer applies a 2×2 pooling window with a stride of 2×2, averaging the values within each window. It smooths out the feature maps and reduces their size, preparing the data for the fully connected layers.
- Flatten Layer: This layer converts the 3D output from the final pooling layer into a 1D vector, bridging the gap between the convolutional layers and the fully connected layers. This transformation is essential for feeding the data into the subsequent layers.
- Fully Connected Layer 1: This layer, with 256 neurons, takes the flattened vector and combines the extracted features into higher-level representations. A ReLU activation function is used to introduce non-linearity, allowing the model to capture complex patterns.
- Fully Connected Layer 2: This layer has 128 neurons and further refines the feature combinations, reducing the dimensionality while maintaining essential information. ReLU activation continues to enable the learning of non-linear relationships.
- Fully Connected Layer 3: Containing 64 neurons, this layer further condenses the information, focusing on the most crucial features. It prepares the data for the final classification step by narrowing down the learned patterns.

- Output Layer: The output layer consists of neurons equal to the number of classes in the task, using softmax activation to produce a probability distribution. The class with the highest probability is selected as the final prediction.

The proposed CNN architecture for classifying DR stages consists of 1,740,645 parameters and starts with an input layer for 224×224 RGB images. It includes three convolutional layers with 32, 64, and 128 filters, progressively capturing more complex features. Max pooling layers follow the first two convolutions, reducing spatial dimensions, while an average pooling layer after the third convolution condenses feature maps. The flattened output is passed through fully connected layers with 256, 128, and 5 neurons. ReLU activation in the fully connected layers captures non-linear patterns, and a softmax-activated output layer classifies images into one of the five DR stages. This architecture offers an efficient solution for multi-stage DR classification.

TABLE I. PARAMETERS OF THE PROPOSED CNN

Layer Name	Type	Customized	Parameters
Input Layer	Input	224×224×3	-
Conv Layer 1	Convolutional	Filters=32, Kernel size=3×3, Stride=1×1	928
Max Pooling 1	Pooling	Pool size=2×2, Stride=2×2	-
Conv Layer 2	Convolutional	Filters=64, Kernel size=3×3, Stride=1×1	18,496
Max Pooling 2	Pooling	Pool size=2×2, Stride=2×2	-
Conv Layer 3	Convolutional	Filters=128, Kernel size=3×3, Stride=1×1	73,856
Average Pooling	Pooling	Pool size=2×2, Stride=2×2	-
Flatten Layer	Flatten	-	-
Fully Connected Layer 1	Fully Connected	Neurons=256	1,605,888
Fully Connected Layer 2	Fully Connected	Neurons=128	32,896
Fully Connected Layer 3	Fully Connected	Neurons=64	8,256
Output Layer	Fully Connected	Neurons=5	325

IV. EXPERIMENTS

This section outlines the experiments conducted, detailing the experimental setup and the configurations used throughout the process.

A. Datasets

The APTOS 2019 dataset [21] is a trusted resource for graded DR severity, and is widely used in clinical practice and research across the Asia-Pacific region. It has been instrumental in developing and evaluating computer-aided diagnostic systems that use retinal images to classify DR severity. The dataset classifies DR into five stages: No DR (0), Mild (1), Moderate (2), Severe (3), and Proliferative DR (PrDR) (4). It includes 3,662 images, and when combined with the DDR dataset, a total of 16,170 images are used for training, validation, and testing. The distribution of stages is 1,014 images for No DR, 999 for Mild, 684 for Moderate, 429 for Severe, and 1,207 for PrDR.

B. Experimental Setup

Approximately 70% of the DR image dataset [21] was used for training and 30% for testing. Five-fold cross-validation (k=5) ensured accurate performance evaluation. The model was trained using binary cross-entropy loss and the Adam optimizer at a 0.0001 learning rate over 50 epochs with a batch size of 32. Python and Keras (TensorFlow-based) were run on an Intel Core i7 CPU with 16 GB RAM and a CUDA-compatible GPU. Multiple trials were used to optimize hyperparameters to balance accuracy and efficiency.

C. Preprocessing

Retinal images were converted to RGB, batched in groups of 32, resized to 224×224 pixels, and scaled. These steps were implemented to improve computational efficiency and enhance feature detection for DR classification.

D. Performance Metrics

Accuracy (*Acc*) [22] was used to evaluate training efficiency, representing the percentage of correct predictions in the test set. Along with precision (*PPV*), sensitivity (*Sen*) [23], and F1-score (*F*1) [24], these metrics provide a comprehensive view of the model's performance, highlighting its robustness. Precision measures the correctness of positive predictions, while sensitivity assesses the identification of actual positives. The F1 balances precision and recall to offer a single effectiveness measure.

$$Acc = \frac{TP+TN}{TP+FP+FN+TN} \quad (8)$$

$$PPV = \frac{TP}{TP+FP} \quad (9)$$

$$Sen = \frac{TP}{TP+FN} \quad (10)$$

$$F1 = 2 \cdot \frac{PPV \cdot Sen}{PPV + Sen} \quad (11)$$

V. RESULTS AND DISCUSSION

This section presents the experimental results, evaluating the proposed combined customized CNN architecture and CLAHE for multi-stage DR classification. The proposed approach was also compared with several previous works. Table II illustrates the experimental results of the proposed CNN architecture alone. Table III presents the experimental results of the combined customized CNN architecture and CLAHE. Figure 3 compares the performance of the custom CNN architecture alone with that of the customized CNN architecture with CLAHE for multi-stage DR classification. The results highlight the significant improvements achieved by incorporating CLAHE into the CNN architecture. The custom CNN alone achieved an overall accuracy of 97.69%, with stage-specific accuracy ranging from 97.5% to 97.85%. In contrast, the combined approach of customized CNN and CLAHE demonstrated a substantial boost in performance, achieving an overall accuracy of 99.69%, with all stage-specific accuracies consistently above 99.6%. This improvement is particularly notable across all stages, showing the effectiveness of CLAHE in enhancing retinal image features, allowing the CNN model to better detect and classify the different stages of DR.

TABLE II. EXPERIMENTAL RESULTS OF THE CUSTOMIZED CNN ARCHITECTURE

DR Stage	*PPV* (%)	*Sen* (%)	*F1* (%)	*Acc* (%)
0	97.5	97.65	97.57	97.6
1	97.4	97.55	97.47	97.5
2	97.7	97.85	97.77	97.8
3	97.6	97.75	97.67	97.7
4	97.8	97.9	97.85	97.85
Overall	**97.6**	**97.74**	**97.67**	**97.69**

TABLE III. EXPERIMENTAL RESULTS OF THE COMBINED CUSTOMIZED CNN ARCHITECTURE AND CLAHE

DR Stage	*PPV* (%)	*Sen* (%)	*F1* (%)	*Acc* (%)
0	99.65	99.7	99.67	99.68
1	99.61	99.66	99.63	99.64
2	99.75	99.77	99.76	99.76
3	99.63	99.65	99.64	99.64
4	99.71	99.73	99.72	99.72
Overall	**99.67**	**99.7**	**99.68**	**99.69**

The proposed approach demonstrated superior performance compared to several state-of-the-art techniques for DR classification. As shown in Table IV, the customized CNN model and CLAHE achieved the highest accuracy of 99.69%, surpassing existing models such as ResNet-152, ResNet-50, and NASNetLarge. The ResNet-152 model in [3] achieved an accuracy of 99.41%, which is slightly lower than the proposed method. Similarly, the ResNet-50 model combined with the enhanced ABC algorithm in [4] reached 98.71%, while NASNetLarge with GSO in [5] obtained 99.26%. Although ShuffleNet-v2 with Median Filtering (MF) in [7] achieved a high accuracy of 99.58%, it is still slightly outperformed by the proposed method. The AnLoG model in [6], which employs LoG for edge detection and PCA for feature reduction, achieved a lower accuracy of 97.29%. Additionally, the CLAHE-enhanced VGGNet and CNN models in [8] and [9] achieved 97.6% and 71.85%, respectively, highlighting the effectiveness of the proposed customized CNN and CLAHE for image enhancement. These results demonstrate that the proposed method not only improves accuracy but also provides a more efficient solution for DR detection by leveraging customized CNN and CLAHE techniques.

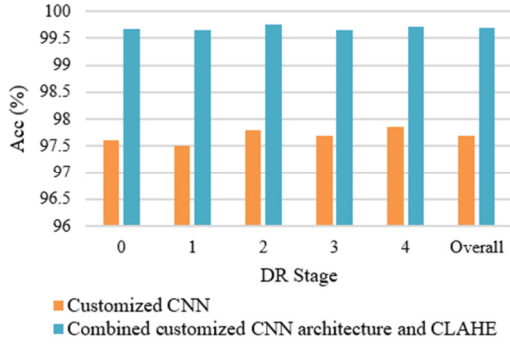

Fig. 3. Comparison of experimental results.

TABLE IV. ACCURACY COMPARISON OF THE PROPOSED WITH STATE-OF-THE-ART MODELS

Study	Year	Number of stages	Image enhancement techniques	Model	Acc (%)
[3]	2022	5	BWO	ResNet-152	99.41
[4]	2023	5	Data augmentation, enhanced Artificial Bee Colony (ABC) algorithm	ResNet-50	98.71
[5]	2023	5	MF for noise removal, GSO for parameter tuning	NASNetLarge	99.26
[6]	2024	4	Laplacian of Gaussian (LoG) for edge detection, Principal Component Analysis (PCA) for feature reduction	AnLoG	97.29
[7]	2024	4	MF for noise reduction	ShuffleNet-v2	99.58
[8]	2024	4	CLAHE	VGGNet	97.6
[9]	2024	5	CLAHE	CNN	71.85
Our Proposed		**5**	**-**	**Customized CNN**	**97.69**
			Customized CLAHE	**Customized CNN**	**99.69**

VI. CONCLUSION

DR is a leading cause of vision loss, making early detection crucial. The key challenge addressed in this study is to achieve high performance in DR detection, while ensuring scalability and computational efficiency, as previous models, despite their accuracy, faced limitations in these areas. This study proposed a customized CNN architecture combined with CLAHE to enhance image quality and improve detection accuracy for multi-stage DR classification. The proposed method achieved an impressive accuracy of 99.69%, outperforming several state-of-the-art models, such as ResNet-152, ResNet-50, and NASNetLarge, which achieved accuracies between 98.71% and 99.58%. The use of customized CNN and CLAHE was crucial in enhancing detection performance. The proposed approach offers an efficient and accurate solution for automated DR classification, supporting early diagnosis and treatment. Future work could involve further optimization of the model and its application to other medical image analysis tasks.

REFERENCES

[1] R. Khandekar, J. A. Lawatii, A. J. Mohammed, and A. A. Raisi, "Diabetic retinopathy in Oman: a hospital based study," *British Journal of Ophthalmology*, vol. 87, no. 9, pp. 1061–1064, Sep. 2003, https://doi.org/10.1136/bjo.87.9.1061.

[2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Aug. 1998, https://doi.org/10.1109/5.726791.

[3] U. Bhimavarapu and G. Battineni, "Deep Learning for the Detection and Classification of Diabetic Retinopathy with an Improved Activation Function," *Healthcare*, vol. 11, no. 1, Jan. 2023, Art. no. 97, https://doi.org/10.3390/healthcare11010097.

[4] U. Bhimavarapu, N. Chintalapudi, and G. Battineni, "Automatic Detection and Classification of Diabetic Retinopathy Using the Improved Pooling Function in the Convolution Neural Network,"

Diagnostics, vol. 13, no. 15, p. 2606, Aug. 2023, https://doi.org/10.3390/diagnostics13152606.

[5] R. Ramesh and S. Sathiamoorthy, "A Deep Learning Grading Classification of Diabetic Retinopathy on Retinal Fundus Images with Bio-inspired Optimization," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11248–11252, Aug. 2023, https://doi.org/10.48084/etasr.6033.

[6] M. D. Ramasamy *et al.*, "A novel Adaptive Neural Network-Based Laplacian of Gaussian (AnLoG) classification algorithm for detecting diabetic retinopathy with colour retinal fundus images," *Neural Computing and Applications*, vol. 36, no. 7, pp. 3513–3524, Mar. 2024, https://doi.org/10.1007/s00521-023-09324-z.

[7] C. Nithyeswari and G. Karthikeyan, "An Effective Heuristic Optimizer with Deep Learning-assisted Diabetic Retinopathy Diagnosis on Retinal Fundus Images," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14308–14312, Jun. 2024, https://doi.org/10.48084/etasr.7004.

[8] A. Jabbar *et al.*, "Correction: Deep Transfer Learning-Based Automated Diabetic Retinopathy Detection Using Retinal Fundus Images in Remote Areas," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, Jun. 2024, Art. no. 145, https://doi.org/10.1007/s44196-024-00557-x.

[9] A. M. Mutawa, K. Al-Sabti, S. Raizada, and S. Sruthi, "A Deep Learning Model for Detecting Diabetic Retinopathy Stages with Discrete Wavelet Transform," *Applied Sciences*, vol. 14, no. 11, May 2024, Art. no. 4428, https://doi.org/10.3390/app14114428.

[10] S. I. S. M. Shazuli and A. Saravanan, "Manta Ray Foraging Optimizer with Deep Learning-based Fundus Image Retrieval and Classification for Diabetic Retinopathy Grading," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11661–11666, Oct. 2023, https://doi.org/10.48084/etasr.6226.

[11] V. T. H. Tuyet, N. T. Binh, and D. T. Tin, "Improving the Curvelet Saliency and Deep Convolutional Neural Networks for Diabetic Retinopathy Classification in Fundus Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8204–8209, Feb. 2022, https://doi.org/10.48084/etasr.4679.

[12] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: speed and effectiveness," in *Proceedings of the First Conference on Visualization in Biomedical Computing*, Atlanta, GA, USA, 1990, pp. 337–345, https://doi.org/10.1109/VBC.1990.109340.

[13] M. A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2015.

[14] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*. O'Reilly Media, Inc., 2017.

[15] Y. Zhu, Q. Ouyang, and Y. Mao, "A deep convolutional neural network approach to single-particle recognition in cryo-electron microscopy," *BMC Bioinformatics*, vol. 18, no. 1, Jul. 2017, Art. no. 348, https://doi.org/10.1186/s12859-017-1757-y.

[16] P. R. R. Kumar and M. Prabhakar, "An Integrated Approach for Lossless Image Compression Using CLAHE, Two-Channel Encoding and Adaptive Arithmetic Coding," *SN Computer Science*, vol. 5, no. 5, Jun. 2024, Art. no. 523, https://doi.org/10.1007/s42979-024-02866-6.

[17] S. Chakraverti, P. Agarwal, H. S. Pattanayak, S. P. S. Chauhan, A. K. Chakraverti, and M. Kumar, "De-noising the image using DBST-LCM-CLAHE: A deep learning approach," *Multimedia Tools and Applications*, vol. 83, no. 4, pp. 11017–11042, Jan. 2024, https://doi.org/10.1007/s11042-023-16016-2.

[18] J. Prakash and K. B. Vinoth, "An ensemble approach for classification of diabetic retinopathy in fundus image," *Multimedia Tools and Applications*, May 2024, https://doi.org/10.1007/s11042-024-19353-y.

[19] S. P. Singh, P. Gupta, and R. Dung, "Diabetic retinopathy detection by fundus images using fine tuned deep learning model," *Multimedia Tools and Applications*, Jun. 2024, https://doi.org/10.1007/s11042-024-19687-7.

[20] N. Gullipalli, V. B. K. L. Aruna, V. Gampala, and B. Maram, "Diabetic retinopathy detection with fundus images based on deep model enabled chronological rat swarm optimization," *Multimedia Tools and Applications*, vol. 83, no. 30, pp. 75407–75435, Sep. 2024, https://doi.org/10.1007/s11042-024-19241-5.

[21] S. R. Rath, "Diabetic Retinopathy 224x224 (2019 Data)." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-2019-data.

[22] W. J. Krzanowski and D. J. Hand, *ROC Curves for Continuous Data*. Chapman and Hall/CRC, 2009.

[23] A. G. Lalkhen and A. McCluskey, "Clinical tests: sensitivity and specificity," *Continuing Education in Anaesthesia Critical Care & Pain*, vol. 8, no. 6, pp. 221–223, Dec. 2008, https://doi.org/10.1093/bjaceaccp/mkn041.

[24] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2022.