IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 November 2024, accepted 9 December 2024, date of publication 17 December 2024,
date of current version 30 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3519361

RESEARCH ARTICLE

Diabetic Retinopathy Classification Using Hybrid Color-Based CLAHE and Blood Vessel in Deep Convolution Neural Network

AMMAR JAWAD KADHIM[1], HADI SEYEDARABI[1], REZA AFROUZIAN[2], AND FADHIL SAHIB HASAN[3]

[1]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666-16471, Iran
[2]Miyaneh Faculty of Engineering, University of Tabriz, Miyaneh 51666-16471, Iran
[3]Electrical Engineering Department, Mustansiriyah University, Baghdad 10052, Iraq

Corresponding author: Reza Afrouzian (afrouzian@tabrizu.ac.ir)

ABSTRACT The most widespread illness of the diabetic eye that causes missing eye vision is diabetic retinopathy (DR), which requires disclosure soon to prevent the vision loss of the sick. In this study, two features are extracted from retina images with for multiclass DR classification, which include color-based Blood Vessel (BV) segmentation and color-based Contrast-Limited-Adaptive-Histogram-Equalization-Top-Hat (CLAH-TH) segmentation. These features are integrated to enhance the accuracy of classification and detection of DR. Variant models, especially VGG19 and InceptionV3, are trained using a transfer learning approach on the proposed extracted features for DR grading. The data augmentation strategy is employed to improve the accuracy and performance of the proposed method by balancing the dataset and aligning the number of images in each class. Experimental results demonstrate that the proposed method outperforms contemporary CNN models when utilizing the suggested features. The best results obtained from experiments on the Kaggle DR database using the pretrained VGG19 model include an accuracy of 96.7%, a sensitivity of 0.971, and a specificity of 0.981.

INDEX TERMS Diabetic retinopathy, feature extraction, pretrained VGG19, pretrained InceptionV3, contrast-limited-adaptive-histogram-equalization, multiclass classification, data augmentation.

I. INTRODUCTION

Diabetic retinopathy (DR) is a prevalent condition among individuals with diabetes and is a significant contributor to vision impairment. Elevated blood glucose levels increase blood viscosity, resulting in fluid accumulation in the tissues surrounding the retina, which ultimately leads to vision loss [1]. According to the International Diabetes Federation report, there were 537 million cases of diabetes worldwide in 2021 and that number is expected to reach about 700 million by 2045 [2]. In the advanced stages of DR, treatment becomes progressively more complex. Accordingly, early diagnosis of DR is crucial for facilitating effective treatment.

A retinal image can be generated by a monocular camera that provides a projection of the fundus onto a 2D plane rapidly and cost-effectively. In the progression of the disease, various features appear in the retinal image, such as blood vessels, hemorrhages, exudates, and neovascularization. These elements constitute the main components of DR progression and are considered as the main lesion features in the fundus image [3]. The fundus image can be enhanced by correcting the illuminations and enhancing the contrast using effective enhancement techniques. The most effective techniques are adaptive-histogram-equalization (AHE) and contrast-limited adaptive-histogram-equalization (CLAHE) [4]. These techniques are utilized to enhance local image contrast as well as noise reduction. In [5], the CLAHE feature is applied to the green channel of the retinal image and used to enhance the blood vessels (BVs). Other advancements to the CLAHE and enhancement techniques, such as edge-aware Laplacian filtering and logarithmic correction, are introduced in [6] and [7], respectively. With advancements in technology

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.

and access to large datasets, modern high-performance techniques based on deep learning (DL) technology have emerged which employed for the automatic analysis and enhancement of image-related data [8], [9], [10].

Transfer learning is a machine learning technique where a model developed for a particular task is reused as a starting point for a model in another task [13]. In this method, instead of training a model from scratch, transfer learning uses the knowledge gained from a pre-trained model that has already learned useful features from a large dataset. This method makes use of the fact that while subsequent layers of deep learning models are increasingly task-specific, the early layers of these models typically learn general features that are relevant to DR classification. This can significantly reduce training time and improve performance, especially in scenarios with small datasets.

Comprehensive research efforts have been proposed to grade DR in various scenarios. These studies employ various feature extraction and classification methods to achieve high accuracy in grading [11], [12]. In all studies, the DR can be classified as either a binary or a multiclass classification. In multiclass classification, DR is categorized into five levels: normal, mild, moderate, severe and proliferative. In contrast, binary classification categorizes DR into only two classes: DR or normal. This method emphasizes multi-class classification and employs DL techniques.

In [13], pre-trained CNN models, including AlexNet, VggNet, GoogleNet, and ResNet are tuned using transfer learning (TL) techniques for DR grading. The EyePACS dataset, containing 36126 images, is utilized for training and evaluation. Experiments results reveal that the VggNet achieves the highest detection accuracy at 95.68%. this study disregards the advantages of image enhancement and trains the CNN directly on fundus images without any pre-filtering. In [14], the authors employed a pretrained CNN to detect DR in the APTOS 2019 and an IDRiD databases. In this work, the authors integrate handcrafted feature with the CNN and optimized them for enhanced performance. The model has achieved an accuracy of 90.07% in classifying data into five categories. This integration has resulted in increased model complexity and computational demands, as well as additional challenges in tuning. In another study, a method utilizing a pre-trained DenseNet CNN model with a transfer learning technique was employed as a classifier on the APTOS 2019 dataset, achieving an accuracy of 96.1% in classifying images to five categories [15]. In similar works, the pre-trained VGG19 model was utilized with the APTOS2019 dataset, comprising of 3662 retinal images, for both training and evaluation. This approach achieved an accuracy of 95.4% in classifying images into five categories. Notably, this study did not employ sophisticated feature extraction approaches; instead, basic filtering and preprocessing techniques were utilized [16]. In addition, in [5], a fine-tuned VGG19 model was employed as a classifier. Various features are extracted, and their combination is applied to the CNN classifier for DR grading. Another work in [17], presents a method for diagnosing DR using deep learning models. Two deep learning models were proposed: the CNN512 model, which classifies the entire image, and the YOLOv3 model, which detects and localizes DR lesions. The fusion of these models achieved an accuracy of 89%

Cao et al. proposed an enhanced residual network (ResNet) model that incorporates attention mechanisms to improve feature extraction and classification accuracy. This method has achieved an accuracy of 91.3% on the EyePACS dataset [18]. Farag et al. also proposed a method for diabetic retinopathy (DR) grading using the DenseNet architecture coupled with a convolutional block attention module. The proposed method enhances feature extraction and achieves an 82% accuracy in classifying the EyePACS dataset into five categories [19]. Albahli et al. explored the design and comparison of several deep learning models for detecting the severity of diabetic retinopathy and segmenting different types of disease patterns. they employed various pre-processing techniques, such as brightness, color, and contrast enhancement, to improve image quality before applying pre-trained models like ResNet50, VGG16, and VGG19 for classification. among them, ResNet50 achieves the best performance. The results demonstrate the effectiveness of image pre-processing and the use of deep learning models in accurately detecting and classifying diabetic retinopathy [20]. Islam et al. utilized a pre-trained Xception CNN model and applied CLAHE for image preprocessing to enhance diabetic retinopathy (DR) detection. Supervised contrastive learning was employed to improve feature representation, resulting in a test accuracy of 98.36% for binary classification of DR and 84.364% for five-stage grading [21]. In another work, Kale and sharma introduced an ensemble deep learning model that successfully identifies five different severity levels of diabetic retinopathy. The ensemble approach integrates multiple deep learning models to improve detection accuracy and robustness [22]. Jian et al. proposed Triple-DRNet, a triple cascade convolutional neural network designed for grading diabetic retinopathy using fundus images. An accuracy of 92.08% was achieved using the APTOS 2019 dataset [23]. Beevi proposed a novel method for classifying diabetic retinopathy (DR) severity using a hybrid optimization-enabled deep learning model. This model merged SqueezeNet and Deep Convolutional Neural Network (DCNN) with optimization techniques [24]. Also, Momeni Pour et al. proposed a method for automatically detecting and monitoring DR using Efficient and CLAHE [36].

Reguant et al. investigated how CNNs such as Inception-v3, ResNet50, InceptionResNet50, and Xception utilized inherent image features for grading DR, using EyePACS and DIARETDB1 datasets [37]. In other work, Jabbar et al. employed VGGNet with transfer learning and data augmentation techniques to diagnose DR from fundus images. It addressed data insufficiency and demonstrated superior accuracy by combining extracted and handcrafted

features [38]. Also, Chilukoti et al. proposed efficient ensemble methods for DR classification, using pre-trained models with transfer learning, CLAHE for image enhancement, and a three-layer classifier with dropout and ReLU activation to enhance feature extraction and generalization. It achieved state-of-the-art quadratic weighted kappa (QWK) scores on Eyepacs, Aptos, and Messidor datasets [39].

Additionally, the authors of publications [41] through [45] provided further research on systems for DR that use improved lesion feature versions and end-to-end deep learning techniques. These methods eliminate the need for distinct steps or manual feature extraction. In [41], the authors employed a multi scale end-to-end network (by EfficientNets) for grading DR. In addition, it used attention module to enrich the features and focus on the crucial regions of the lesions. This network utilized Messidor1 and IDRID datasets for training and evaluation. DR grade for Messidor-1 dataset is categorized into two classes (referable, non-referable) and for IDRID dataset is categorized into five classes. Based on the experiment results of the article, the best accuracy for Messidor1 (for two classes) and IDRID (for 5 classes) is 97.5% and 70.1%, respectively.

The authors in [42] proposed a modified U-Net architecture, termed RMCA U-net, which enhances the segmentation performance of hard exudates by incorporating residual connections, channel attentions and multi-scale feature extraction. Also, Fu et al. developed a novel end-to-end architecture for diabetic macular edema (DME) grading that combined ResNet50 with channel attention (SENet) for feature extraction [43]. This work utilized Messidor dataset and grades DME to three classes with accuracy of 97%. In [44], the authors used a U-Net architecture enhanced with a probabilistic bubble technique to segment the optic disc in abnormal fundus images. The results show improved segmentation compared to traditional techniques, especially in images with abnormalities. The authors in [45] introduced a new approach that used geometric features of blood vessels to infer the location of the fovea. By analyzing the orientation and distribution of blood vessels, this method effectively identified the foveal area.

Despite the numerous methods presented for the diagnosis and grading of DR, challenges such as generalization, variability in fundus image quality, data imbalance, and complexity of disease manifestations persist. Additionally, some works have a complex structure and include several parts that have limitations in their training and how they communicate with each other to produce the desired output. On the other hand, understanding the workings and decision-making processes of these systems is challenging.

This study attempts to provide solutions to some of remaining challenges and increase the accuracy in classifying the severity of diabetic retinopathy. The contributions of this article are summarized in the following paragraphs:

- To enhance the robustness of grading systems against variations in clarity, sharpness, and lighting of fundus photographs, this study employs color-based blood vessel (BV) segmentation and color-based contrast limited adaptive histogram equalization top-hat (CLAH-TH) methods.
- This study employs the Kaggle Diabetic Retinopathy (DR) dataset, with Table 1 displaying the image distribution in different categories and highlighting the imbalance in the dataset. To address this imbalance and improve the model generalization, the data augmentation technique is utilized.
- Microaneurysms, hemorrhages, and exudates are just a few of the symptoms that can be subtle and challenging to identify in the early stages of DR. Therefore, the lesions caused by diabetic retinopathy can be diverse in appearance and involve different areas in the retina. As a result, the proposed methods for diagnosing and classifying this type of disease should have high generalizability. In the proposed method, at first, it has been tried to identify the appropriate features extracted from retinal images that are most effective in the classification of diabetic retinopathy. Then the selected features were applied to the CNN-based model in different modes (individually and simultaneously) to evaluate their performance. After performing various experiments, the best result has been obtained by the simultaneous employing of features related to Color-Based CLAHE segmentation and Blood Vessel segmentation.
- To find the best CNN backbone network for DR classification, several well-known architectures were tested. VGG19 and InceptionV3 demonstrated superior performance in DR grading. These architectures were further refined by adjusting the top layers to achieve the highest accuracy and lowest loss.

The remaining of the paper can be shortened as follows: Section II provides a complete discussion of the methodology and its materials. The experimental results and analysis of the proposed method are presented in detail in section III. finally, in section IV, the conclusion is described.

II. METHODOLOGY

The Kaggle Diabetic Retinopathy (DR) Dataset is used in this work, with 35,126 fundus images that are tagged as right and left eyes for training and on the website (https://www.kaggle.com/c/diabetic-retinopathy-detection/data). It has five output categories: zero (no DR), one (mild), two (moderate), three (severe), and four (proliferative). Table 1 illustrates the image number for each category in the Kaggle DR dataset. Fig. 1 presents sample fundus images for each class, taken from the Kaggle DR Dataset. In the no DR class, labeled as zero, no lesions are present. the mild class labeled as one, shows slight micro-aneurysms and soft exudates. The moderate class features additional micro-aneurysms, soft exudates and minor hemorrhages. The server class exhibits in increased number

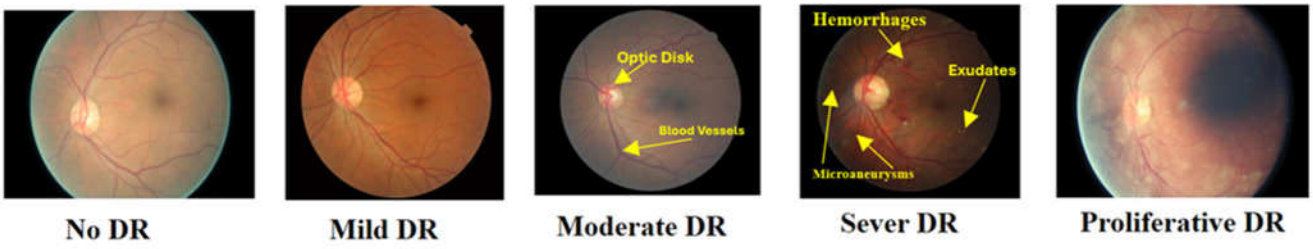


FIGURE 1. Sample fundus images for each class taken from the Kaggle DR dataset.

of hemorrhages, while the proliferative class characterized by further hemorrhages and the presence of hard exudates.

As shown in previous works, selecting appropriate features and applying them to CNNs, instead of end-to-end learning, can lead to increased interpretability, computational efficiency, and improved system accuracy. On the other hand, the selection of one feature alone may not be able to provide all the required information in the image for classification of DR. Therefore, in this article, various studies were conducted to select the appropriate features and their simultaneous use for the DR classification, and the highest accuracy has been obtained by the simultaneous employing of features related to Color-Based CLAHE segmentation and Blood Vessel segmentation.

Fig. 2 presents the block diagram of the proposed method for grading DR. initially the method applies color-based Blood Vessel (BV) segmentation and color-based Contrast Limited Adaptive Histogram Equalization (CLAHE) Top-Hat (CLAH-TH) to the input retinal image. The outputs of these methods are then combined and fed into the proposed network based on CNN. These features provide important, effectively aiding in the training and learning of patterns for different classes of DR. The combined features block is responsible for concatenating the two features and feeding them to the proposed CNN-based network. A detailed description of each block in the Fig. 2 is provided in the following subsections.

As mentioned, The Kaggle Diabetic Retinopathy (DR) Dataset is used in this work, with 35,126 fundus images. Table 1 illustrates the image number for each category in the Kaggle DR dataset. as be shown, the number of images for each category differ from each other. To balance the dataset and enhance model generalization, some strategies like oversampling, under-sampling, or employing synthetic data generation are used. This manuscript employs data augmentation techniques to compensate for the low number of samples in some categories. The data augmentation technique addresses the imbalance in the number of class images by augmenting the retinal images using functions such as left-right flipping, rotation, translation [35]. Fig. 3 illustrates an example of data augmentation on the retinal image, where the original image is transformed by rotation and flip functions. Images, according to the number of images in each class, undergo this data augmentation strategy and then passed to the feature extraction to obtain the required features. subsequently, the proposed network, based on VGG19 or InceptionV3 models, is employed DR classification.

TABLE 1. Number distribution of images in the Kaggle DR dataset.

Classes	DR Diagnosis	Images Number
Zero	Normal	25810
One	Mild DR	2443
Two	Moderate DR	5292
Three	Sever DR	873
Four	Proliferative DR	708

A. COLOR-BASED BLOOD VESSEL (BV) SEGMENTATION

In Fig. 4, the block diagram of the proposed color-based blood vessel segmentation method is presented. The input retinal image is in RGB color format, comprising three channels: red, green, and blue. Let's define the input image and its three channels as $I = [I_R, I_G, I_B]$, where I represents the original RGB retinal signal and I_R, I_G and I_B correspond to the red, green and blue channels, respectively. Each channel has an 8-bit representation and is considered a gray-scale image. The following procedures are provided to obtain the color vessel image:

1- We extract the red, green, and blue channels from the original retinal image, denoted as I_i, where $i \in \{R, G, B\}$.

2- each channel is then filtered using Gaussian filtering. This method enhances the grayscale image using the 2D Gaussian kernel function $\phi(x, y)$. The output of Gaussian filtering for the ith channel is calculated as [25]:

$$f_i(x, y) = \phi(x, y) \circledast I_i(x, y), \quad i \in \{R, G, B\} \tag{1}$$

where $\circledast$ is the convolution operator and $\phi(x, y)$ is the two-dimensional Gaussian filter, expressed as [26]:

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{2}$$

where σ^2 is the variance of the two-dimensional Gaussian filter.

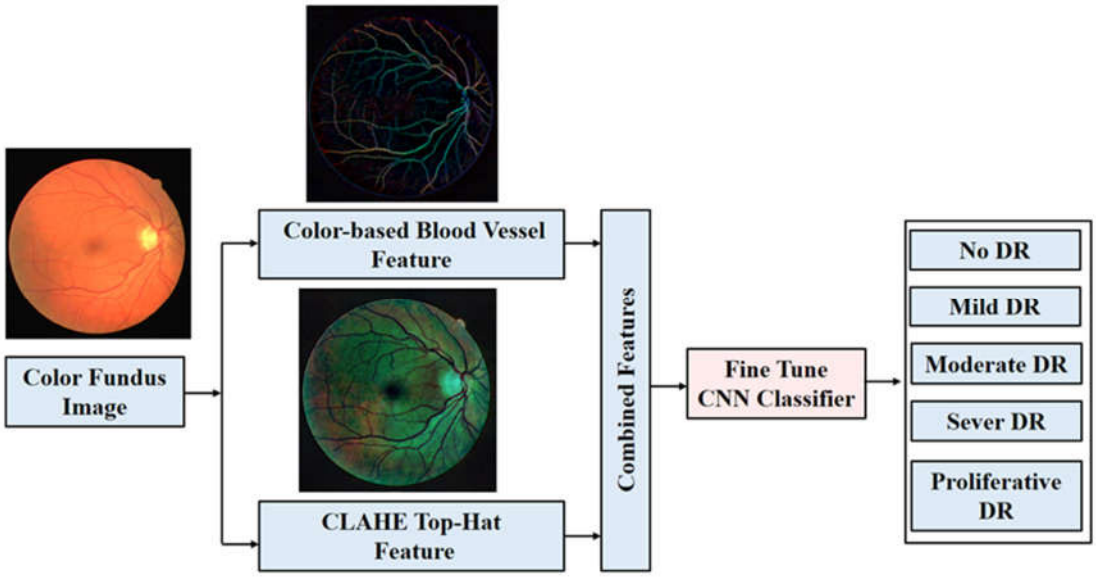
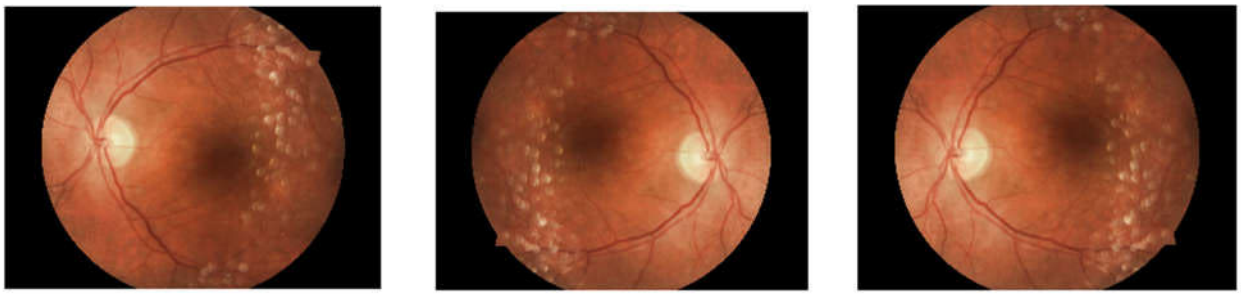


FIGURE 2. Schematic diagram of the suggested DR detection system.



(a) Original image. (b) Rotated image by 180°. (c) Flipped image.

FIGURE 3. Example of data augmentation.

The vessel segmentation function proposed in [1] is applied to each channel to obtain the *i*th vessel image. Then, the three channels are combined to create the color vessel image.

Fig. 5 illustrates the block diagram of the vessel segmentation algorithm applied to each channel of the image. Two types of segmentations, thick and thin, are combined using morphological operations. Both types of segmentation use optimized-top-hat technique, as described in [25]. The method of the vessel segmentation function can be summarized as follows:

1) SEGMENTATION OF THICK VESSELS

1- the optimized-top-hat technique is applied to the *i*th output of the Gaussian filter to improve the bright objects that are the center of interest against a dark background in an image. The optimized top-hat output can be expressed as follows [25]:

$$t_{o,i} = f_i^c - (f_i^c \circ S_o) \blacksquare S_c, \quad i \in \{R, G, B\} \tag{3}$$

where S_o is the structural elements of the opening ($\circ$) function and S_c represents the structural element of the closing ($\blacksquare$) function. f_i^c represents the image complement of the *i*th image (f_i), expressed as:

$$f_i^c(x, y) = \max(U) - f_i(x, y) \tag{4}$$

where U represents the generic set of potential values for one sample. The radii are set as eight and 16 pixels for the opening ($\circ$) and closing ($\blacksquare$) functions, respectively.

2- homomorphic filtering (HF) is applied to the *i*th output of the optimized-top-hat function. HF enhances the grayscale image using multiple frequency-domain filters to regulate the lighting intensity and reflection of the image [25], [27]. The *i*th output of the HF process, at position (x, y), can be written as:

$$t_{hf,i}(x, y) = e^{\left(\mathcal{F}^{-1}(L_{o,i}(u,v)).\psi(u,v)\right)} \tag{5}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the two-dimensional DFT and IDFT, respectively. $L_{o,i}(u, v)$ is the frequency domain of the

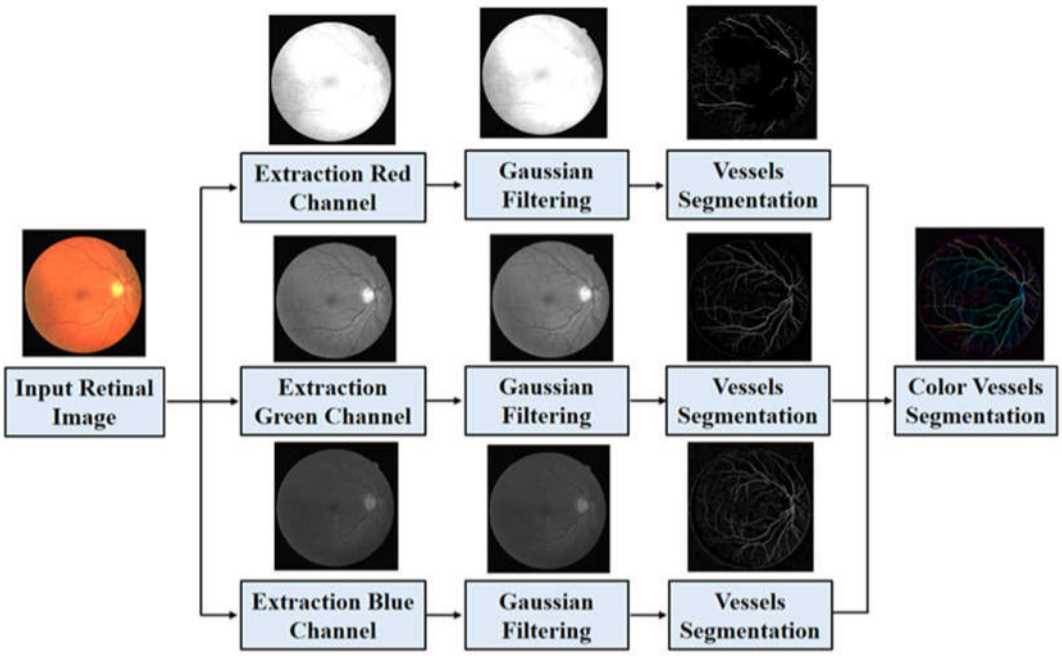


FIGURE 4. Color-based blood vessel segmentation block diagram.

log scale of the $t_{o,i}(u, v)$ signal, which is expressed as: $L_{o,i}(u, v) = \mathcal{F}\left(logt_{o,i}(u, v)\right)$. The function $\psi(u, v)$ is the frequency domain of Gaussian high-pass response, which can be written as [25]:

$$\psi(u, v) = 1 - \exp(-\frac{(u^2 + v^2)}{2\sigma_{HF}^2}) \tag{6}$$

where σ_{HF}^2 is the criterion that determines the diffusion of the Gaussian curve.

3- A median filter is applied to remove the undesired noise produced by the HF process, which is attributed to the low value of σ_{HF}^2. The output of the median filter at the ith channel is expressed as:

$$t'_{hf,i}(x, y) = Me\left(t_{hf,i}(n, m), (n, m) \in \mathcal{R}\right) \tag{7}$$

where Me represents the median value of the pixels. The neighborhood size is a 2×2 matrix. After that, the enhanced top-hat is applied to the output of the median filter to improve the thick vessels. The radii of 32 and 86 are used for opening and closing operations, respectively.

2) DISTRIBUTION OF THIN VESSELS

The first two steps of thin vessel segmentation consist of the same steps in thick vessel segmentation. The only difference is in the selection of opened and closed disc-shaped skeleton elements with a radius of 4 and 20 pixels, respectively [25].

4- in the second step, two-dimensional matched filtering (MF2) is applied to the output of the HF process, $t_{o,i}(x, y)$,. The purpose of MF2 is to enhance the image regions through a specific distribution. The matched filter kernel, which follows a Gaussian-shaped curve, is defined as follows [25] and [28]:

$$h_{MF}(u, v) = exp\left(\frac{-u^2}{2\sigma_{MF}^2}\right), \text{f or } |y| \leq \mathcal{L}/2 \tag{8}$$

where $\mathcal{L}$ is the multi-partite element length and σ is a kernel density profile spread (herein $\mathcal{L}$ is 7 and σ is 0.8). The distribution (u, v) is the rotation of (x, y) by an angular step θ that is defined as [25] and [28]:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{9}$$

The output of MF2 is produced by convolving $t_{o,i}(x, y)$ and $h_{MF}(u, v)$ According to:

$$I_{MF2,i}(x, y) = t_{o,i}(x, y) \circledast h_{MF}(u, v), \ i \in \{R, G, B\} \tag{10}$$

where the symbol $\circledast$ denotes the two-dimensional convolutional function.

5- The Minimum-Cross-Entropy-Threshold-Harris-Hookes-Optimization (MCETHHO) technique is utilized on the output of MF2 functions. This technique employs HHO to minimize the statistical criteria for express entropy pass, aiming to obtain the optimal threshold and thereby simplify image intensity levels. Let the **th** be defined as a vector that has different threshold values for the sector signal, and the cross-entropy function based on this is expressed as [25]:

$$CE(\boldsymbol{th}) = \sum_{l=1}^{L} ih(l) \ln(l) - \sum_{l=1}^{nt} H_l \tag{11}$$

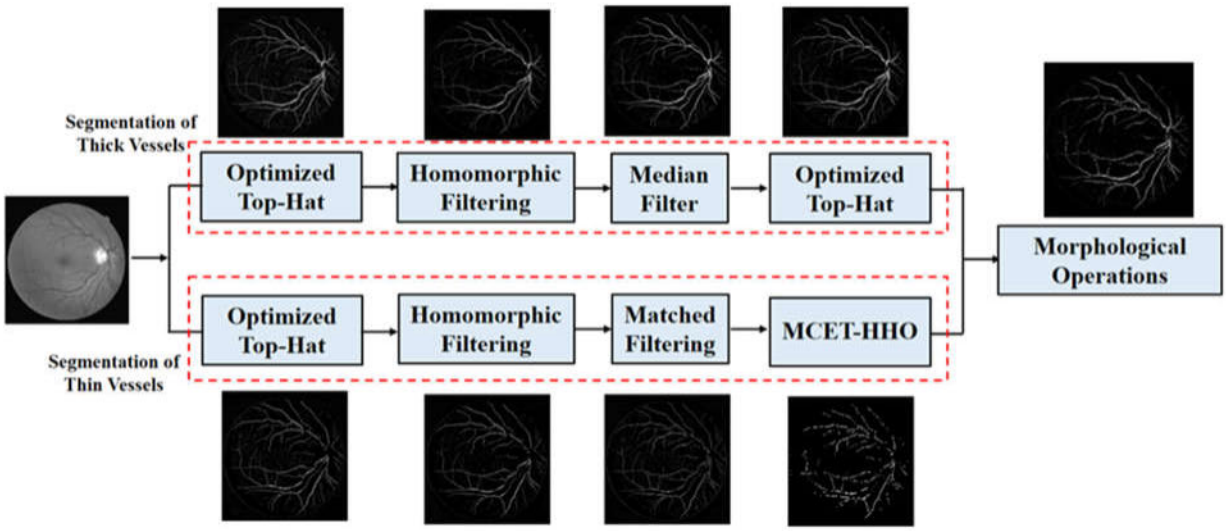


FIGURE 5. The blood vessels segmentation block diagram.

where nt is the number of thresholds, and H_l can be expressed as:

$$H_1 = \sum_{l=1}^{th_1-1} ih(l) \ln(\mu(1, th_1))$$
$$H_k = \sum_{l=1}^{th_k-1} ih(l) \ln(\mu(th_{k-1}, th_k)),$$
$$H_{nt} = \sum_{l=1}^{th_{nt}-1} ih(l) \ln(\mu(th_{nt}, L+1)) \quad 1 < k < nt \quad (12)$$

The end of the iteration in the HHO process can occur when the minimum fitness values are reached, evaluating the optimal threshold in the cross-entropy function. The HHO parameters are set as follows: 250 iterations, 30 falcons for the algorithm, and the four optimal thresholds, $th_l, l = 1, 2, 3, 4$. The ith output image of the MCET-HHO technique is expressed as:

$$I_{MCET-HHO,i}(x, y) = \begin{cases} th_1, & if\ I_{MF2,i}(x, y) \le th_1 \\ th_l, & if\ th_{l-1} < I_{MF2,i}(x, y) < th_l, l = 1, 2, .., nt-1 \\ th_{nt}, & if\ I_{MF2,i}(x, y) > th_{nt} \end{cases} \quad (13)$$

Lastly, the ith output image $I_{MCET-HHO,i}$ is quantized for the thin vessels feature.

3) MERGE THE SEGMENTATIONS

By utilizing morphological image operations after merging the thick and thin segmentations in (1) and (2), enhanced vessel segments are produced for the ith channel.

B. CLAHE TOP-HAT FEATURE

Fig. 6 illustrates the proposed color-based CLAHE Top-Hat segmentation. The input retinal image is represented in a color format with RGB channels. The following procedures are provided to obtain the color CLAHE Top-Hat segmentation: Extracting the red, green, and blue channels from the original retinal image, I_i, $i \in \{R, G, B\}$.

1- each channel is filtered using CLAHE method.
2- then, Top-Hat filtering is applied to their outputs.
3- Finally, the color CLAH-TH segmentation is obtained by merging three images.

Fig. 7 illustrates the comparison of the color retinal image between the original, CLAHE-filtered image and CLHE-TH segmentation.

C. FINE TUNE CNN CLASSIFIER

so far, various architectures have been proposed for convolutional neural networks. Selecting the appropriate network from these options is the initial step in designing the proposed model. For this purpose, various studies were conducted in this paper, ultimately resulting in the selection of the VGG19 and Xception networks as the final models. In the subsequent step, a detailed design was developed for the dense layers to ensure high-accuracy DR classification of the features extracted by these networks.

As illustrated in Fig. 2, the two aforementioned features, color-based BV and CLAHE-TH segmentation, are extracted from the retinal image. After being concatenated, these features are applied to the specified architectures. The DL classification contains two stages: the training and testing processes. during the training phase, the training features are applied to the DL classifier to set the optimal parameters of the system. In the testing phase, the classes are assigned to the testing set or unclassified images. The abnormal appearances in retinal images vary significantly; therefore, they need to be extracted individually from given retinal images.

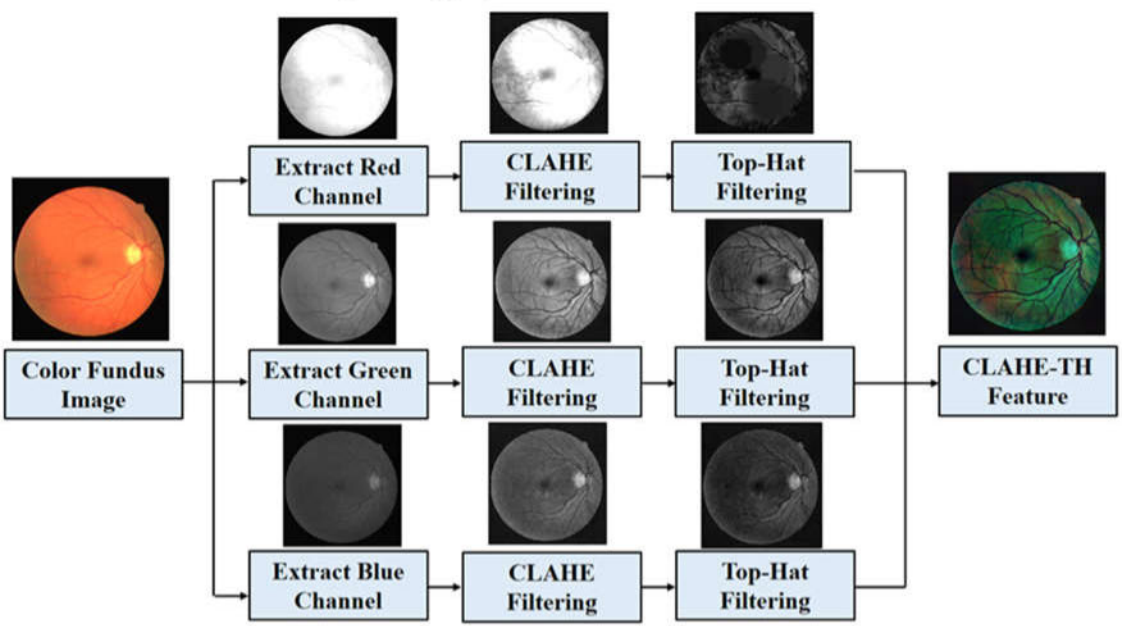


FIGURE 6. The CLAHE-TH segmentation block diagram.

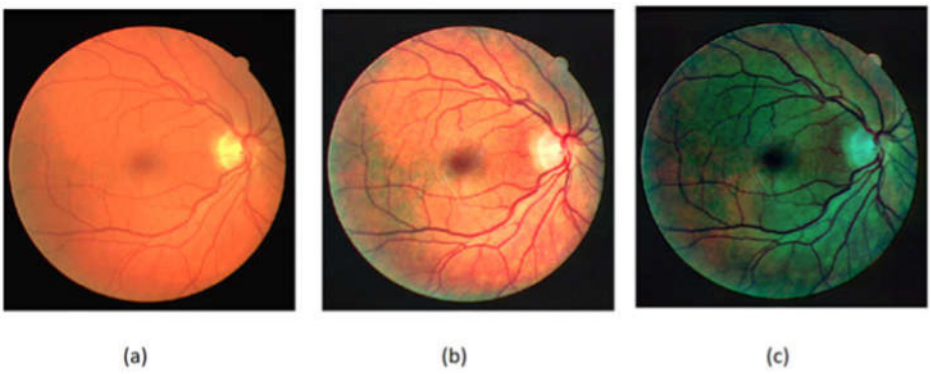


FIGURE 7. The Color retinal image. (a) The original picture. (b) CLAHE filtered picture. (c) CLAHE-TH image.

CNN models have a remarkable track record for the interpretation of images. various architectures have been developed, with a particular focus on transfer learning methodologies. The CNN weights consist of filters, where the dimension of each filter is an $n \times n$ matrix that convolves (slides and multiplies) over the input data. The optimal values for these filters are obtained through the CNN training process, which involves passing the input image through multiple layers, each layer composed of specific neurons that process different aspects of the image, ultimately leading to the final layer that identifies the correct output.

In transfer learning, the features and weights learned from previously trained models are used to train new models with as little data for new tasks. This approach improves processing time during the training stage and enhances the results. Table 2 illustrates the details of standard pre-trained models. The input dimensions represent the size of the input images, while depth indicates the topological depth of the CNN model.

TABLE 2. The models and their details.

Models	Parameters	Depth	Input dimensions
AlexNet [29]	62,378,344	13	227×227
GoogLeNet [30]	5,438,750	27	224× 224
VGG16 [31]	138,357,544	23	224× 224
VGG-19 [32]	143,667,240	26	224× 224
Xception [33]	22,910,480	126	299× 299
InceptionV3 [34]	23,851,784	159	299× 299
Fine-tuned VGG-19 [7]	143,667,240	26	224× 224
Pretrained VGG19	**21,077,057**	**27**	**320×320× 3**
Pretrained InceptionV3	**21,077,057**	**27**	**320×320× 3**

In this study, fine-tuning is utilized to enhance the performance of pre-trained VGG19 and InceptionV3 networks

TABLE 3. The number of images for each class for individual and combined features.

DR Severity	Individual Features	Combined Features	Training images percentage (85%)	Test images percentage (15%)
No DR	1875	3750	3187	562
Mild	1480	2960	2516	444
Moderate	1993	3986	3388	598
Severe	1520	3040	2584	456
Prolific	1760	3520	2992	528
Total	8628	17256	14667	2588

in the classification of diabetic retinopathy images. For this purpose, the weights of the initial layers of the networks are frozen, and training is conducted only on the dense layers and some of the final layers of the mentioned networks. Consequently, according to the information in Table 2, only 21,077,057 parameters are trained, which significantly reduces both learning time and data requirements.

III. EXPERIMENTAL RESULTS

The aim of This study is to classify retinal images into multiple classes and distinguish between no DR images and abnormal ones based on the severity of abnormalities, including mild, moderate, severe, and proliferative diabetic retinopathy (DR). The data set is divided into training and testing sets using a predefined ratio of 85:15. Two features, color-based blood vessel segmentation and color CLAHE-TH, are input into the proposed networks. Additionally, this article proposes two models: one based on the pretrained VGG-19 architecture and the other on the pretrained InceptionV3 architecture. These models are fine-tuned on feature derived from Kagle database images.

The performance of the models is evaluated using criteria such as accuracy, specificity, and sensitivity. Accuracy represents the ratio of correctly classified samples to the total number of samples:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{14}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{15}$$

$$Specificity = \frac{TN}{TN + FP} \tag{16}$$

where TP represents the true positive value, TN the negative value, FP the false positive value, and FN the false negative value.

Table 3 illustrates the number of images for each class for individual and combined features after augmentation. The combined features have doubled the number of images input to the CNN classifier. Additionally, for each network, various hyperparameters are configured. These parameters are fine-tuned through a trial-and-error approach. The optimal hyperparameter values for the networks proposed in this article are presented in Table 4.

The accuracy, sensitivity, and specificity of multi-class classification based on different pretrained VGG19 and InceptionV3 CNN models for individual and combined.

TABLE 4. The optimal parameter values for the proposed networks.

Parameters	Values	parameters	Descriptions
Epochs	95	**Optimizer**	Adam
Learning rate	1e-4	**Metric list**	Accuracy, Specificity, Sensitivity
Image size	320×320× 3	**Loss**	Mean square error
Depth	27		

	No DR	Mild	Moderate	Severe	Prolific
No DR	**545**	**10**	**5**	**2**	**0**
Mild	**10**	**424**	**8**	**2**	**0**
Moderate	**1**	**10**	**578**	**7**	**2**
Severe	**0**	**3**	**7**	**436**	**10**
Prolific	**0**	**3**	**5**	**10**	**510**

FIGURE 8. Classification details of test images using a confusion matric.

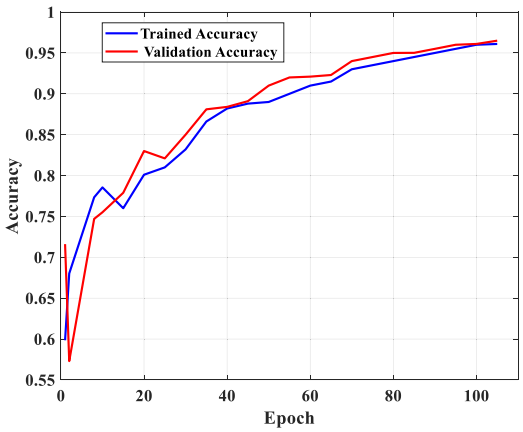


FIGURE 9. Accuracy measure for training and validation process.

As previously mentioned, one of the key innovations of this paper is the examination of various feature extraction and enhancement methods for diabetic retinopathy classification. For this purpose, different methods were analyzed, some of which are presented in Table 5. It is evident that the Color-based BV and Color CLAHE-TH methods exhibit the highest accuracy in DR grading. Furthermore, given their commendable performance, another experiment was conducted where both methods were applied to retinal images. Following integration, they were simultaneously input into the proposed neural networks. As shown in Table 5,

TABLE 5. Comparison of various features extraction and image enhancing methods for DR grading using VGG19 and Inception based networks.

Features	Pretrained VGG19			Pretrained InceptionV3		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
Blood vessels	72.1	0.74	0.721	73.2	0.744	0.734
Exudates	73.2	0.731	0.742	74.3	0.75	0.74
CLAHE	75.3	0.78	0.77	76.2	0.77	0.78
Color based BV	80.8	0.812	0.81	82.2	0.843	0.833
Color CLAHE-TH	87.8	0.883	0.874	85.1	0.871	0.841
Combined feature	95	0.951	0.881	95.6	0.96	0.94
Combined feature after data augmentation	**96.7**	0.971	0.981	96.5	0.97	0.96

TABLE 6. Comparison between the proposed methods and other recent works in grading the severity of DR in retinal images.

Model	Database	Number of classes	Accuracy %	Sensitivity	Specificity
[7]	DRIVE/STARE	2	94.92/95.65	0.71/0.71	**0.98**/**0.98**
[14]	IDRiD	5/3	90.07/96.85	-	-
[17]	DDR/ Kaggle	5	88.6/84.1	-	-
[18]	Kaggle	5	91.3	0.91	0.97
[21]	Kaggle	5	84.3	-	-
[22]	Kaggle	5	87.31	-	-
[23]	Kaggle	5	92.0	-	-
[41]	IDRID/ Messidor1	5/2	70.1/**97.5**	-	-
Pretrained InceptionV3	Kaggle	5	96.5	0.96	0.97
Pretrained VGG19	Kaggle	5	96.7	**0.981**	0.971

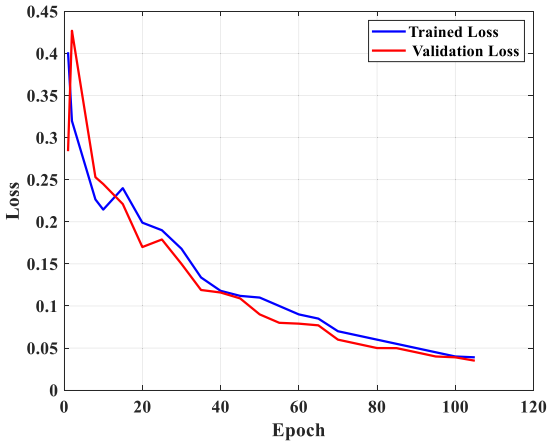


FIGURE 10. Loss measure for training and validation process.

the highest accuracy in diabetic retinopathy image classification is achieved through the simultaneous application of the mentioned features to the proposed network based on VGG19.

For grading diabetic retinopathy, various databases have been utilized in the articles presented in this field. Moreover, these databases differ in terms of the number of classes used to classify the severity of retinopathy images, ranging from two to five classes. Table 6 compares the methods proposed in this article with recent methods. To ensure a fair comparison, the table includes the name of the database used by other works and the number of classes used for classification. As observed, the method presented in this paper has achieved the highest accuracy in classifying Kaggle database images into five defined categories.

Fig. 8 presents the classification details of test images through a confusion matrix. To evaluate the training method of the proposed network in this paper based on the InceptionV3 model, accuracy and loss curves are plotted for the training and validation data during the learning period, in Fig. 9 and Fig. 10 respectively. These images demonstrate that the proposed network is effectively trained through the selection of an appropriate model, the precise design of the compression layer, and the careful adjustments of hyperparameters.

IV. CONCLUSION

Automated classification of DR presents an effective way for tracking and preventing vision loss in a significant number of patients. In this paper, two retinal features are proposed, which are color-based blood vessel (BV) segmentation and color-based constrained adaptive histogram top hat equation (CLAHE). Both techniques improve the low-quality fundus photographs and, hence, grade the DR with sufficient accuracy. Two proposed networks, based on VGG19 and InceptrionV3, are employed to classify retinal images using transfer learning method. Additionally, a data augmentation strategy is used to improve accuracy and performance by balancing the dataset and increasing the number of images The experimental results indicate that the extracted features in this article outperform conventional ones that use the image as input. Also, the proposed pre-trained models detect different grades of DR with higher accuracy.

REFERENCES

[1] H. Safi, S. Safi, A. Hafezi-Moghadam, and H. Ahmadieh, "Early detection of diabetic retinopathy," *Surv. Ophthalmol.*, vol. 63, no. 5, pp. 601–608, Sep. 2018.

[2] Int. Diabetes Fed., Brussels, Belgium. *International Diabetes Federation Diabetes Atlas*. Accessed: Dec. 2021. [Online]. Available: https://www.diabetesatlas.org/en/

[3] A. Malhi, R. Grewal, and H. S. Pannu, "Detection and diabetic retinopathy grading using digital retinal images," *Int. J. Intell. Robot. Appl.*, vol. 7, no. 2, pp. 426–458, Jun. 2023.

[4] Y. Qi, Z. Yang, W. Sun, M. Lou, J. Lian, W. Zhao, X. Deng, and Y. Ma, "A comprehensive overview of image enhancement techniques," *Arch. Comput. Methods Eng.*, vol. 29, no. 1, pp. 583–607, Jan. 2022.

[5] A. Dutta, P. Agarwal, A. Mittal, and S. Khandelwal, "Detecting grades of diabetic retinopathy by extraction of retinal lesions using digital fundus images," *Res. Biomed. Eng.*, vol. 37, no. 4, pp. 641–656, Dec. 2021.

[6] P. Raja Rajeswari Chandni, J. Justin, and R. Vanithamani, "Fundus image enhancement using EAL-CLAHE technique," in *Advances in Data and Information Sciences*. Singapore: Springer, 2022, pp. 613–624.

[7] S. Gross, M. Klein, and D. Schneider, "Segmentation of blood vessel structures in retinal fundus images with logarithmic Gabor filters," *Current Med. Imag. Rev.*, vol. 9, no. 2, pp. 138–144, May 2013.

[8] Q. You, C. Wan, J. Sun, J. Shen, H. Ye, and Q. Yu, "Fundus image enhancement method based on CycleGAN," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 4500–4503.

[9] S. Woo, J. Park, J. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, Jan. 2018, pp. 3–19.

[10] G. Ali, A. Dastgir, M. W. Iqbal, M. Anwar, and M. Faheem, "A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images," *IEEE J. Transl. Eng. Health Med.*, vol. 11, pp. 341–350, 2023.

[11] S. Panchal and M. Kokare, "A comprehensive survey on the detection of diabetic retinopathy," *IETE J. Res.*, vol. 70, no. 1, pp. 236–256, Jan. 2024.

[12] A. M. Dayana and W. R. S. Emmanuel, "A comprehensive review of diabetic retinopathy detection and grading based on deep learning and metaheuristic optimization techniques," *Arch. Comput. Methods Eng.*, vol. 30, no. 7, pp. 4565–4599, Sep. 2023.

[13] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," *Comput. Electr. Eng.*, vol. 72, pp. 274–282, Nov. 2018.

[14] B. Harangi, J. Toth, A. Baran, and A. Hajdu, "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 2699–2702.

[15] K. Swathi, E. S. N. Joshua, B. D. Reddy, and N. T. Rao, "Diabetic retinopathy detection using deep learning," in *Proc. Int. Conf. Advancement Smart, Secure Intell. Comput. (ASSIC)*, Nov. 2022, pp. 1–5.

[16] M. H. Islam and M. Nahiduzzaman, "Severity grading of diabetic retinopathy using deep convolutional neural network," *Int. J. Innov. Sci. Res. Technol.*, vol. 6, p. 7, Feb. 2021.

[17] 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, p. 3704, May 2021.

[18] J. Cao, J. Chen, X. Zhang, Q. Yan, and Y. Zhao, "Attentional mechanisms and improved residual networks for diabetic retinopathy severity classification," *J. Healthcare Eng.*, vol. 2022, pp. 1–10, Mar. 2022.

[19] M. M. Farag, M. Fouad, and A. T. Abdel-Hamid, "Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module," *IEEE Access*, vol. 10, pp. 38299–38308, 2022.

[20] S. Albahli and G. N. Ahmad Hassan Yar, "Automated detection of diabetic retinopathy using custom convolutional neural network," *J. X-Ray Sci. Technol.*, vol. 30, no. 2, pp. 275–291, Mar. 2022.

[21] M. R. Islam, L. F. Abdulrazak, M. Nahiduzzaman, M. O. F. Goni, M. S. Anower, M. Ahsan, J. Haider, and M. Kowalski, "Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105602.

[22] Y. Kale and S. Sharma, "Detection of five severity levels of diabetic retinopathy using ensemble deep learning model," *Multimedia Tools Appl.*, vol. 82, no. 12, pp. 19005–19020, May 2023.

[23] M. Jian, H. Chen, C. Tao, X. Li, and G. Wang, "Triple-DRNet: A triple-cascade convolution neural network for diabetic retinopathy grading using fundus images," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106631.

[24] S. Zulaikha Beevi, "Multi-level severity classification for diabetic retinopathy based on hybrid optimization enabled deep learning," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104736.

[25] O. Ramos-Soto, E. Rodríguez-Esparza, S. E. Balderas-Mata, D. Oliva, A. E. Hassanien, R. K. Meleppat, and R. J. Zawadzki, "An efficient retinal blood vessel segmentation in eye fundus images by using optimized top-hat and homomorphic filtering," *Comput. Methods Programs Biomed.*, vol. 201, Apr. 2021, Art. no. 105949.

[26] R. C. Gonzalez and R. Woods, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[27] C.-N. Fan and F.-Y. Zhang, "Homomorphic filtering based illumination normalization method for face recognition," *Pattern Recognit. Lett.*, vol. 32, no. 10, pp. 1468–1479, Jul. 2011.

[28] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 3, pp. 263–269, Sep. 1989.

[29] F. N. Landola, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size," 2016, *arXiv:1602.07360*.

[30] P. Ballester and R. M. Araujo, "On the performance of GoogleNet and AlexNet applied to sketches," in *Proc. 30th AAAI Conf. Artif. Intell.*, vol. 30, no. 1, 2016.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–14.

[32] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, p. 1, Dec. 2018.

[33] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.

[34] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[35] M. S. Patil, S. Chickerur, C. Abhimalya, A. Naik, N. Kumari, and S. Maurya, "Effective deep learning data augmentation techniques for diabetic retinopathy classification," *Proc. Comput. Sci.*, vol. 218, pp. 1156–1165, Jan. 2023.

[36] A. Momeni Pour, H. Seyedarabi, S. H. Abbasi Jahromi, and A. Javadzadeh, "Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization," *IEEE Access*, vol. 8, pp. 136668–136673, 2020.

[37] R. Reguant, S. Brunak, and S. Saha, "Understanding inherent image features in CNN-based assessment of diabetic retinopathy," *Sci. Rep.*, vol. 11, no. 1, p. 9704, May 2021, doi: 10.1038/s41598-021-89225-0.

[38] M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, "Transfer learning-based model for diabetic retinopathy diagnosis using retinal images," *Brain Sci.*, vol. 12, no. 5, p. 535, Apr. 2022, doi: 10.3390/brainsci12050535.

[39] S. V. Chilukoti, L. Shan, V. S. Tida, A. S. Maida, and X. Hei, "A reliable diabetic retinopathy grading via transfer learning and ensemble learning with quadratic weighted Kappa metric," *BMC Med. Informat. Decis. Making*, vol. 24, no. 1, pp. 1–12, Feb. 2024, doi: 10.1186/s12911-024-02446-x.

[40] Y. Fu, Y. Ju, and D. Zhang, "MSEF-Net: A multi-scale efficientnet fusion for diabetic retinopathy grading," *Biomed. Signal Process. Control*, vol. 98, Dec. 2024, Art. no. 106714, doi: 10.1016/j.bspc.2024.106714.

[41] Y. Fu, G. Zhang, X. Lu, H. Wu, and D. Zhang, "RMCA U-Net: Hard exudates segmentation for retinal fundus images," *Expert Syst. Appl.*, vol. 234, Dec. 2023, Art. no. 120987, doi: 10.1016/j.eswa.2023.120987.

[42] Y. Fu, X. Lu, G. Zhang, Q. Lu, C. Wang, and D. Zhang, "Automatic grading of diabetic macular edema based on end-to-end network," *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 118835, doi: 10.1016/j.eswa.2022.118835.

[43] Y. Fu, J. Chen, J. Li, D. Pan, X. Yue, and Y. Zhu, "Optic disc segmentation by U-Net and probability bubble in abnormal fundus images," *Pattern Recognit.*, vol. 117, Sep. 2021, Art. no. 107971, doi: 10.1016/j.patcog.2021.107971.

[44] Y. Fu, G. Zhang, J. Li, D. Pan, Y. Wang, and D. Zhang, "Fovea localization by blood vessel vector in abnormal fundus images," *Pattern Recognit.*, vol. 129, Sep. 2022, Art. no. 108711, doi: 10.1016/j.patcog.2022.108711.
[45] M. Alharbi and D. Gupta, "Segmentation of diabetic retinopathy images using deep feature fused residual with U-Net," *Alexandria Eng. J.*, vol. 83, pp. 307–325, Nov. 2023, doi: 10.1016/j.aej.2023.10.040.

AMMAR JAWAD KADHIM received the B.Sc. degree in electrical engineering from Mustansiriyah University, Iraq, in 2006, and the M.Sc. degree in telecommunication system engineering from the Belarusian State University of Informatics and Radio-Electronics, Belarus, in 2015. He is currently pursuing the Ph.D. degree with the Faculty of Electrical and Computer Engineering, University of Tabriz, Iran. His research interests include telecommunication systems, image processing, and deep learning.

HADI SEYEDARABI received the B.S. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 1993, the M.S. degree in electrical engineering from the K. N. T. University of Technology, Tehran, Iran, in 1996, and the Ph.D. degree in electrical engineering from the University of Tabriz, in 2006. He is currently a Professor with the Faculty of Electrical and Computer Engineering, University of Tabriz. His research interests include image processing, computer vision, video coding, human–computer interaction, computational brain modeling and simulation, and deep learning applications in image processing, medical signal processing, and computer vision.

REZA AFROUZIAN received the Ph.D. degree in electrical engineering from the University of Tabriz, Tabriz, Iran, in 2015. He is currently an Assistant Professor with Miyaneh Faculty of Engineering, University of Tabriz, Miyaneh, Iran. His research interests include image processing, computer vision, pattern recognition, and machine learning.

FADHIL SAHIB HASAN received the B.Sc. degree in electrical engineering and the M.Sc. degree in electronics and communication engineering from Mustansiriyah University, Baghdad, Iraq, in 2000 and 2003, respectively, and the Ph.D. degree in electronics and communication engineering from Basrah University, Iraq, in 2013. In 2005, he joined the Faculty of Engineering, Mustansiriyah University. He is currently a Professor with Mustansiriyah University. His research interests include wireless communication systems, multicarrier systems, wavelet-based OFDM, MIMO systems, speech and image signal processing, chaotic cryptography, chaotic modulation, FPGA, and Xilinx System Generator-based communication systems.