Retinal Image Enhancement Through Hyperparameter Selection Using RSO for CLAHE to Classify Diabetic Retinopathy

Selvamani Hemamalini[1*] , Visvam Devadoss Ambeth Kumar[2] , Venkatesan Ramachandran[3] , Rene Robin[4]

[1] Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Anna University, Chennai 600123, India
[2] Department of Computer Engineering, Mizoram University, Aizawl 796004, India
[3] Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore 641114, India
[4] Department of Computer Science and Engineering, Sri Sairam Engineering College, Anna University, Chennai 600044, India

Corresponding Author Email: hemamalini@panimalar.ac.in

ABSTRACT

Diabetic retinopathy (DR) is one of the main causes of blindness in diabetes patients. To choose the appropriate course of action for the patient, the severity level of the problem must be verified to avoid vision loss. Deep learning models are being utilized to classify the stage of DR at which the patient is right now. However, the input image to be fed to the Convolutional Neural Network (CNN) must be of good quality to facilitate subsequent tasks such as image segmentation, feature extraction, and classification. Most image enhancement techniques rely on Histogram modification, especially contrast-limited adaptive histogram equalization (CLAHE), to perform local contrast enhancement in order to avoid the drawbacks of existing methods. With Improper selection of hyperparameters such as the number of tiles and clipping limit, an image's quality decreases and becomes inappropriate for further processing. The uniqueness of Random Search Optimization (RSO) lies in its ability to dynamically adapt the search process based on the behavior of virtual rat swarms, allowing for effective exploration of the parameter space while balancing exploration and exploitation. RSO's advantage over existing methods lies in its ability to effectively handle high-dimensional and nonlinear search spaces, potentially leading to the discovery of optimal hyperparameter configurations that enhance retinal images and improve diabetic retinopathy classification accuracy. This study proposes an optimized clipping limit selection using the Rat Swarm Optimization algorithm for contrast-limited adaptive histogram equalization to enhance the retinal fundus images. The enhanced images are further segmented using Otsu thresholding for classification using CNN. The proposed model was evaluated and tested on the fundus datasets MESSIDOR and IDRiD. The result shows that the proposed optimized clipping limit selection using the Rat Swarm Optimization algorithm model outperforms with 7.33 MSE, 44.51 PNSR, and 0.89 SSIM values when compared to other methods. Classification accuracy was also improved by enhancing the fundus images.

1. INTRODUCTION

Diabetic retinopathy is a diabetes-related eye disease that affects the retina, the light-sensitive tissue at the back of the eye. It's one of the leading causes of blindness in adults. The condition develops when high levels of blood sugar cause damage to the blood vessels in the retina. When diseases are discovered early, they are easier to treat and are more likely to be successful. Insulin deficiency leads to diabetes, a disease that causes an increase in blood glucose levels [1]. Around the world, it affects 425 million adults [2]. The retina's blood vessels enlarge and leak fluid and blood [3]. This is known as diabetic retinopathy (DR). In an advanced stage of the disease, blindness is possible. 2.6% of all blindness incidents are caused by DR [4]. A higher incidence of diabetes-related complications is seen in those with diabetes who have had the condition for an extended period of time. To prevent blindness, diabetics must have regular retinal examinations to detect and treat diabetic retinopathy (DR) at an early stage [5]. On a retinal scan, numerous types of lesions can be seen that indicate the presence of DR. Few examples of lesions may be aneurysms (MA), hematomas (HM), and soft and hard exudates (EX) [6, 7].

As a result of the weakening walls of blood vessels, microaneurysms (MA) are the early sign of DR, which show as minute red spherical spots. The size is less than 125 microns and the edges are angular. Also discovered six different types of MA. To categorize the various types of MA, conventional imaging using fluorescein and adaptive optics scanning laser ophthalmoscopy (AOSLO) was also used [8]. Haemorrhages (HM) in the retina are large, irregularly bordered spots that occur on the retina. In terms of HM, there are two types: flame

(the most visible HM) and blot deeper HM. Plasma leakage causes bright yellow patches on the retina to look like hard exudates. They have distinct borders and are found in the outer layers of the retina [9]. Cotton wool-like exudates (also known as soft exudates) show as white spots on the retina due to nerve fiber enlargement. Typically, it has an oval or circular shape. Bright lesions are soft and hard exudates (EX), whereas red lesions are HM. When these lesions are present, there are five different levels of DR according to the severity of the damage: minimal damage; mild damage; moderate damage; moderate damage; severe damage; and proliferative damage.

Diabetic retinopathy (DR) affects diabetics in two ways: as a moderate form called NPDR, or as a more severe form called Proliferative Diabetic Retinopathy (PDR). Exudates are the first symptom of DR and suggest a milder form of the disease, known as Non-Proliferative Diabetic Retinopathy (NPDR). However, the retina generates new blood vessels that severely impair vision in patients with NPDR [10, 11], although the patient initially experiences just blurred vision. Blood clots and blots may form in the retina as a result of aberrant blood vessels that leak or bleeding [12]. DR is almost always caused by damage to the retina's blood vessel network. Blood vessels are completely blocked in the advanced stages of PDR, resulting in lesions in the blood vessels. Haemorrhage and aneurysm are the two most prominent lesions that have occurred. Microaneurysms, which appear as red dots in the fundus, are the first visible sign of DR. The challenge lies in efficiently navigating the high-dimensional hyperparameter space and finding settings that can generalize well accross diverse retinal images in order to improve the diagnostic performance of classifying diabetic retinopathy.

1.1 Detection of DR

Currently, diagnosis is made manually by an ophthalmologist, which is more prone to misdiagnosis and more time-consuming than automatic procedures. An automated DR screening device is needed that can detect the condition quickly and reliably. In terms of time and money, automated DR detection approaches are superior to manual methods [13]. A new unsupervised method has been used to construct a system with great accuracy. Deep-learning-based techniques are used to detect and segment the fovea at the pixel level in the retinal picture. A Convolutional Neural Network (CNN) is a type of deep neural composed of multiple layers of neurons that is used in deep learning. For every neuron in a layer, it is linked to every neuron in every other layer. The field of image categorization offers many uses for CNN.

1.2 Present limits to the diagnosis of DR

CNN arts have been successful in the diagnosis of DR, but they still have significant limits, which are detailed in the following paragraphs:
1) The image quality could be significantly degraded due to poor hyperparameter selection.
2) The current CNN models for DR diagnosis focus only on grading DRs without recognizing the exact location of each DR lesion in the fundus picture. This is a limitation.

1.3 Contribution

Three steps are used for classification of DR images, where

pre-processing step includes gray-scale conversion, image contrast, and resizing the images. During this step, the clipping limit for limiting the contrast in using the CLAHE algorithm is defined by using the Rat Swarm Optimization algorithm to enhance the image quality. Then, the binary conversion, overlay and morphological operations along with small pixel removal processes are presented in the segmentation process. Then, the classification is carried out using CNN model. By considering this pre-processing and segmentation model, all images without leaving single images are used from datasets for classification.

1.4 Association of the paper

- The analysis of current methods for classifying the DR from fundus pictures is covered in Section 2.
- A brief explanation of the proposed model is given in Section 3.
- The validation of the projected model with existing procedures is presented in Section 4.
- Finally, the scientific contribution of the work is depicted in Section 5.

2. LITERATURE REVIEW

The morphological method for accurate segmentation was addressed by Wang et al. [14]. It is used in deep learning to consider the co-occurrence matrix on the texture. The entire procedure works in stages to pinpoint the precise position of the infection in the retina. Aside from that, segment-based CNN (S-CNN) uses two hidden layers to address diagnosis defects and classify classification-based thresholds and standards. Using an exemplary bag of topographies, Leeza and Farooq [15] created a diabetic retinopathy detection model. The dictionary technique used in both preprocessing and postprocessing techniques has been used to improve the detection of diabetics. The pathological learning outline is used mainly to combine it. Hemanth et al. [16] projected an upgraded DR model for the identification and classification using the deep neural network. Contrast development strategies are limited in this study, as image processing is required to equalise histograms.

Shankar et al. [17] have proposed an automated deep learning-based model for the identification and classification algorithm for fundus diabetic retinopathy (DR) pictures. Preprocessing, segmentation, and classification are all components of the new approach. A pre-processing stage is used to remove any unwanted noise from the edges. Following histogram-based segmentation, the most valuable parts of the image are extracted. The SDL model is then applied for the classification of the DR fundus images according to varying degrees of severity. The Messidor DR Dataset is used to support the validity of the SDL model that is being presented. Based on the results of the tests, the SDL model described here is superior to the previous models in terms of classification accuracy.

A two-layer CNN is used to improve the evaluation of layers, which reduces the number of untruthful during classification [18]. The precise location of the unhealthy area of the retina is determined by this successive process. S-CNN, on the other hand, uses twofold hidden layers to discriminate among the threshold and normalised positions based on classification to manage the diagnostic fault.

To carry out effective detection, Bhimavarapu and Battineni [19] combined fuzzy logic approaches with digital image processing. Particle swarm optimization was used to segment the digital fundus images and find microaneurysms in the retina. Clustering high-similarity data into clusters was used to combine membership functions in particle swarm optimization. The DIARETDB0 dataset was used for model testing and images were segmented using probability-based clustering techniques (PBPSO). The effectiveness of existing fuzzy models and the probability discrete particle swarm optimization method were compared. The findings demonstrated that the proposed particle swarm optimization algorithm has a 99.9% accuracy rate in identifying DR at an early stage.

Supervised contrastive learning (SCL) has been presented by Islam et al. [20] as a way to overcome issues including low margins that lead to incorrect findings, sensitivity to noisy data, and hyperparameter fluctuations. In this study, using the "APTOS 2019 Blindness Detection" dataset, fundus images (FI) were used to identify DR and its severity stages using the SCL method, a two-stage training method with supervised contrastive loss function. The "Messidor-2" dataset was also used to test the model performance. The pretrained Xception CNN model was used as the encoder, and transfer learning and contrast limited adaptive histogram equalization (CLAHE) were employed. The embedding space, which is composed of a 128D space, was represented using t-SNE to understand the model's SCL in a 2D space. The proposed model correctly recognized DR (Binary classification) using the APTOS 2019 dataset with 98.36% accuracy and an AUC score of 98.50%, as well as with 84.364% accuracy and an AUC score of 93.819% for five levels of classification. The performance of this model has also been assessed using different assessment measures (precision, recall, F1 score), in addition to APTOS 2019 and Messidor 2. The suggested method also beat classic CNN without SCL and other cutting-edge algorithms in detecting DR [21], thus supporting its efficacy.

3. PROPOSED APPROACH

The objective of this research is to classify DR fundus photos with the highest possible detection rate utilizing effective preprocessing and segmentation algorithms. DR affects a large number of people, so it is important to classify them into distinct stages. Figure 1 shows the whole process, and the processes are described below.

3.1 Dataset description

To determine DR, the benchmark MESSIDOR dataset is used. This article has about 1200 color fundus pictures with the appropriate annotation. This collection contains four different kinds of images. Based on the presence or absence of micro-aneurysms and hemorrhages. In an image without symptoms of injury, a healthy retina can be observed. The figure shows stage 1 with a few microaneurysms. Small blood vessels and hemorrhages are classified as stage 2, while larger blood vessels, such as those in the brain, are classified as stage 3. This study also used the Indian Diabetic Retinopathy Image Data Set (IDRiD) [22]. This publicly accessible dataset includes 516 fundus photos with labels at a 50-degree FOV.

3.2 Pre-processing

Pre-processing photos is important to eliminate noise, improve image characteristics, and ensure image consistency [23]. Low contrast and non-illumination in retinal colour fundus images are caused by anatomical variables including the 3-D concave form of the eye fundus, opaque media therein, wide-angle lenses on cameras, pupil size fluctuation, sensor array geometry, and movement during image collection. The pre-processing of retinal fundus pictures is crucial as a result. Preprocessing techniques try to increase the likelihood that a disease will be detected using computer-aided segmentation of retinal images and visual assessment. In the first stage of the procedure, color fundus images from two datasets are scaled to 128 by 128 pixels. To accelerate processing, images are converted to grayscale. The original image is shown in grayscale and in scaled form in Figure 2.
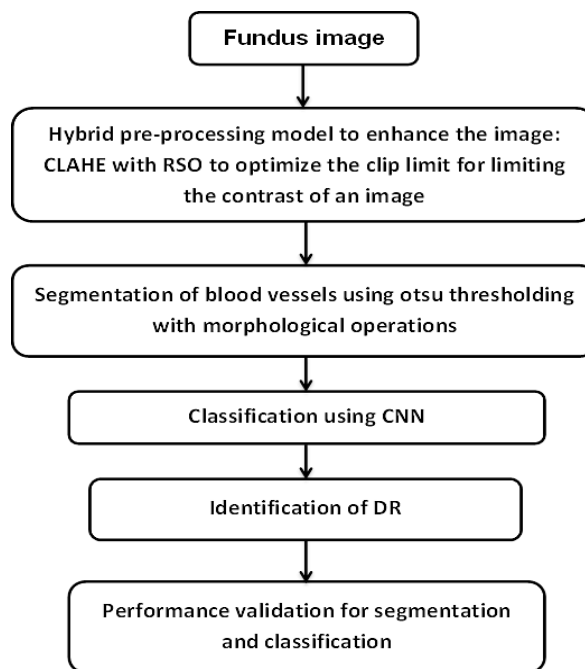

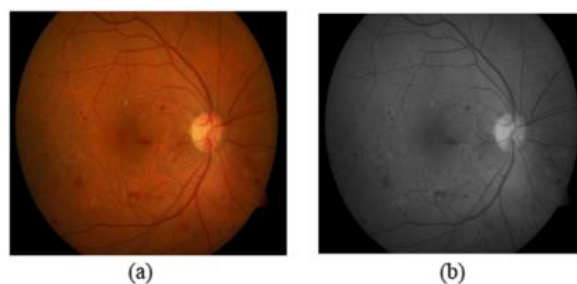**Figure 1.** Overall workflow of the projected scheme

Figure 2. a) Resized RGB image, b) Gray image

These pictures consist of a fundus image in the colors red, green, and blue. The green channel of the RGB image is used for preprocessing due to the strong contrast between the blood vessels and the background and the best contrast between the optic disc and the retinal tissue. In the red channel, the choroid veins are very distinct. Although the green channel has decreased contrast, the retinal vessels can still be clearly seen. The grayscale image will be used to localize the optical disc.

The blue channel is inadequate for identification due to noise and lack of retinal morphological data.

After grayscale conversion, histogram equalization is used to improve contrast and uniformity. A grayscale image is transformed into a histogram equalised image using image enhancement techniques. The focus here is on highlighting the finest details of the image. Histogram-equalized images make it possible to see microaneurysm, optical disc, and hard exudates in more detail. It is a type of Adaptive Histogram Equalization (AHE) algorithm known as CLAHE. Clip limit and tile count settings are used to solve the problem of AHE's over amplification. The image is divided into MN local tiles by CLAHE. The histogram is generated independently for each tile. The selection of these parameters can significantly impact the quality of image enhancement and subsequent classification performance. However, the paper may lack explicit details on how these parameters are chosen or optimized. To compute the histogram, we first need to compute the average number of pixels (N_s) in each section using Eq. (1).

$$N_s = (N_X \times N_Y)/N_G \tag{1}$$

The average total number of pixels, N_G, is determined by multiplying N_X by two and N_Y by two. The number of pixels in the X dimension is represented by N_X, and the number of pixels in the Y dimension is represented by N_Y. The Rat Swarm Optimization algorithm determines the best clip limit solution because it maintains a good balance between investigation and manipulation [24]. The clip limit (N_A) of CLAHE is utilized to improve the image. Rat Swarm Optimization algorithm is utilized to determine the clip limit in this study.

3.2.1 An optimization algorithm with a mathematical model

The chasing and fighting behaviors of the rat are covered in this section. Next, a description of the Rat Swarm Optimization (RSO) algorithm follows.

3.2.2 Hunting the prey

Rats are often sociable animals that act in socially agonistic ways to chase their prey in packs. It is assumed that the finest search agent knows exactly where the prey is. The other search agencies' rankings with respect to the top search agent at the moment are subject to change. Eq. (2) is defined to determine the best search agent:

$$\vec{T} = U\,\vec{T}_i(x) + K\left(\vec{T}_r(x) - \vec{T}_i(x)\right) \tag{2}$$

The best-optimal clipping limit is $\vec{T}_r(x)$, the location of the pixels is $\vec{T}_i(x)$, and the placement of the rats is defined.

The parameter U is calculated using Eq. (3) and Eq. (4).

$$U = L - x \times \left(\frac{L}{Max_{iteration}}\right) \tag{3}$$

where,

$$x = 0,1,..,Max_{iteration}$$
$$K = 2\,rand\,() \tag{4}$$

The random numbers L and K fall between (1, 5) and (0, 2), respectively. Parameters U and K become increasingly crucial

as iterations progress in order to maximize both investigation and exploitation.

3.2.3 Battling a prey

To statistically represent the conflict between rats and their prey, Eq. (5) has been proposed.

$$\vec{T}_i(x+1) = \left|\vec{T}_r(x) - \vec{T}\right| \tag{5}$$

In this case, the rat's new updated position is defined as $\vec{T}_i(x+1)$. The other search agent's positions are updated in accordance with the best solution's position when the best solution is saved. The rat (G, H) can shift positions in order to follow its prey (G, H). Eqs. (3) and (4) state that changing the parameters will result in a different number of places surrounding the current position. However, this notion can be expanded further in an n-dimensional context. Exploration and exploitation are guaranteed by the adjusted values of the parameters U and K. The recommended RSO approach uses the fewest number of operators while maintaining the best result. The choice of RSO as an optimization algorithm for a specific problem or task can depend on various factors and considerations.

Algorithm 1 presents the pseudo-code of the RSO algorithm.

Algorithm 1: Proposed RSO algorithm.
Input: the rats population $T_i(i = 1,2,...,n)$
Output: the optimal search agent for clipping limit
1.   *Initialize the parameters U, K and L*
2.   *Calculate the fitness value of each search agent*
3. $T_r \leftarrow$ *the best search agent*
4.   *While $(x < Max_{iteration})$ do*
5.   *for each search agent do*
6.   *Update the position of current search agent by Eq. (5)*
7.   *End for*
8.   *Update parameters U, K and L*
9.   *Check if there is any search agent which goes beyond the given search space and then adjust it*
10.   Calculate the fitness of each search agent
11.   *Update T_r if there is a better solution than previous optimal solution*
12. $x \leftarrow x + 1$
13.   *End while*
14.   *Return T_r*
15.   *End procedure*

Then, by finding the optimal solution for the clip limit, Eq. (6) defines how to clip the histogram (N_{CL}).

$$N_{CL} = N_A \times N_{optim} \tag{6}$$

In Eq. (6), N_A is the clip limit that is identified by rat optimization and the novelty lies in this process and N_{optim} is the normalized clip limit between 0 and 1. The clip limit for the height of the histogram is then applied to each tile.

$$H_i = \begin{cases} N_{CL} & if\ N_i \geq N_{CL} \\ N_i & else \end{cases} \quad i = 1,2,...,1-1 \tag{7}$$

where, N represents the number of grey levels, L represents the height of the histogram of the i-th tile and H represents the height of the histogram of the i-th tile in Eq. (7). Eq. (8) can be used to calculate the total number of pixels cut.

$$N_C = (N_X \times N_Y) - \sum_{i=0}^{L-1} H_i \qquad (8)$$

This is the number of pixels that have been cropped in Eq. (8). We must also redistribute the clipped pixels after computing N_C. Nonuniformly or uniformly, the pixels can be rearranged. We can use Eq. (9) to figure out how many pixels need to be redistributed.

$$N_R = N_C/L \qquad (9)$$

The number of pixels to be redistributed (N_R) is mentioned in Eq. (9). Eq. (10) is used to normalise the histogram after it has been cut.

$$H_i = \begin{cases} N_{CL} & if \ N_i + N_R \geq N_{CL} \\ N_i + N_R & else \end{cases} i = 1,2,\dots,1-1 \qquad (10)$$

Eq. (8) and (9) are used to get the total number of pixels that are not evenly distributed across the image. Eq. (10) is repeated until all pixels have been redistributed. Eq. (11) is being used to express the contextual region's cumulative histogram.

$$C_i = \frac{1}{(N_x \times N_y)} \sum_{j=0}^{i} H_j \qquad (11)$$

Only after all computations are performed does the histogram of the contextual region match a probability distribution that offers a predetermined brightness and visual quality. Assume that P(x, y) has the value T and that C_1, C_2, C_3, and C_4 are the four centre points of the neighbour tiles that belong to P(x, y). The weighted total is calculated over these four circumstances. The independent tiles are combined using bilinear interpolation, the artifacts between the independent tiles are removed, and the new value of T can be calculated using Eq. (12).

$$T' = (1-y)\big((1-x) \times C_1(T) + x \times C_2(T)\big) \\ + y(1-x) \times C_3(T) + x \times C_4(T)) \qquad (12)$$

After these procedures have been completed, the enhanced image can be obtained. On the basis of the existence of specific components, a picture might be categorized as dark, bright, or low-contrast. The histogram (0-255) and the grey level image (0-255) both have an axis at the number of pixels in the image with a horizontal range of 0 to 1. During picture capture, noticeable freckled noise might be picked up due to the sensor's surroundings. This problem is solved using a noise reduction method. The average filter and background subtraction techniques are used in this process. Figure 3 shows the result of the enhanced image after applying the proposed algorithm.
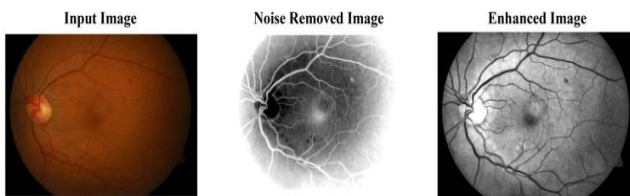

Figure 3. Output of pre-processed images

3.3 Segmentation

3.3.1 Otsu thresholding

In some sections of the image, the Wiener filter softens noise. As a result, values in the grayscale range ranging from white to black can be found. It is necessary to apply Threshold in order to get rid of these grey level values. The image has been subjected to the Otsu thresholding method to accomplish this. An introduction to threshold values is provided using the Otsu approach. For the purpose of classifying images, a single threshold value is used in this investigation. C0 and C1 are the two classes of the image. Eq. (13) is used to calculate the probability distribution. It involves identifying and delineating different anatomical structures and lesions within the retinal images. Here's a more detailed description of the segmentation process, along with potential novel contributions and comparisons with existing techniques: Thresholding techniques separate regions of interest from the background based on pixel intensity values. Common thresholding methods include global thresholding, adaptive thresholding, and Otsu's method.

$$P_i = h(i)/N \qquad (13)$$

The total number of pixels in the image is represented by the pixel numbers $h(i)$ and N. The weight of each cluster is derived using Eq. (14) by adding up the grey level possibilities in C0 and C1;

$$W_0 = \sum_{i=0}^{t-1} P_i \quad W_1 == \sum_{i=t}^{L-1} P_i \qquad (14)$$

The average values of the clusters are intended as Eq. (15).

$$\mu_0 = \sum_{i=0}^{t-1} \left(\frac{i.P_i}{W_0}\right) \mu_1 = \sum_{i=t}^{L-1} \left(\frac{i.P_i}{W_1}\right) \qquad (15)$$

Eq. (16) is used to get the average density of the image as a whole.

$$\mu_t = \mu_0 W_o + \mu_1 W_1 \qquad (16)$$

The alterations of classes are intended as Eq. (17).

$$\sigma_0 = W_0(\mu_0 - \mu_t)^2 \ \sigma_1 = W_1(\mu_1 - \mu_t)^2 \qquad (17)$$

The objective function for the Otsu process is intended using Eq. (18).

$$f(t) = \sigma_0 + \sigma_1 \qquad (18)$$

Lastly, the threshold value exploits the objective function is originating.

3.3.2 Morphological opening

Due to the thresholding approach, the blood vessels appear white, while the background is dark. However, there is still some noise in the image. Also, these noises fall under the category of "white." It has to be taken out of the picture. For this purpose, structural differences between blood arteries and noise are investigated. As you can see in this retinal image, there are many small pixel groups that make up noise, but they

have an uneven structure. These traits provide the morphological openness needed to remove the noise. The size of the structural element is 4. The morphological opening method is carried out using Eq. (19).

$$X°B = (X \ominus B) \oplus B \qquad (19)$$

X is a picture and B is a structural element. Only noise is removed in the morphological opening process because blood vessels and noise do not come into contact.

3.3.3 Circle-removal

Due to the limitations of the fundus camera, an image takes on a circular shape. It is important to remove this circle. To do this, the Hough transform is used to determine the shape and then is eliminated from the image. Figure 4 is an example output of the application.
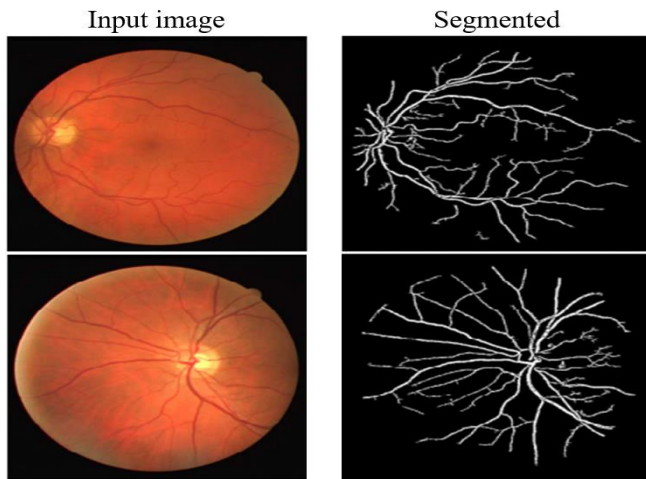

Figure 4. Sample output of segmentation

3.4 Classification

DL is categorized as a subset of ML techniques for Artificial Intelligence (AI) platforms. CNN gives the model the ability to use traditional technology to develop, explore, and simplify laborious paradigms. The ability to interpret data representations at multiple layers is made possible by multiple levels of abstraction. CNN are supervised DL dependent approaches that produce the best advances in image processing. The combination of convolutional, fully connected (FC) and grouping is the three main CNN layers. Segmented images are fed into the CNN method to determine the various classes of two datasets. Two training phases were implemented for CNN. The induced input images occurred during the feed forward phase. Each neuron's input vector dot product and parameter vector are computed concurrently. Then the achieved result is determined. A loss function has been used to compare the applied results and estimate the error rate based on the error. Each parameter gradient calculation is done using the chain rule, leading to the final tailoring of the entire parameters. Maximum processing rounds will be completed after the performance.

CNN Architecture Details:
Number of layers: Specify the number of convolutional layers, pooling layers, and fully connected layers in the CNN architecture.

Filter sizes: Provide information on the size of filters used in convolutional layers.
Activation functions: Mention the types of activation functions used in different layers (e.g., ReLU, sigmoid, tanh).
Pooling operations: Specify the pooling operations (e.g., max pooling, average pooling) and their sizes.
Dropout: If dropout regularization is used, specify the dropout rates and the layers where dropout is applied.
Output layer: Describe the configuration of the output layer, including the number of nodes and the activation function used.

3.4.1 Initializing the weight

The perfect initial weights accelerate network convergence. Here, a number of initialization methods can be established for the weights. Second, it can be seen through the analysis of using various initializers that the normal distribution through the initializer He produces the best performance.

3.4.2 Activation functions

Deep networks typically use an activation function or non-linear operator for the following convolutions. The concept is improved by the existence of the function. A well-liked convolution in DNN that uses rectified linear units (ReLU) accelerates training. Eq. (20) explains how ReLU calculates negative rates that are heading toward zero.

$$ReLU(u) = \begin{cases} u \; if \; u \geq 0 \\ 0 \; if \; u < 0 \end{cases} \qquad (20)$$

Eq. (21) incorporates the leaky ReLU and provides it in relation to ReLU. Due to the fact that it is not active, the function allows for a small, non-zero gradient [25].

$$leaky_{relu}(u) = \begin{cases} x \; if \; u \geq 0 \\ ax \; if \; u < 0 \end{cases} \qquad (21)$$

where, $\alpha = 0.3$. Exponential Linear Units (ELUs) are now thought to increase training speed and classifier accuracy. However, because of the low calculation cost, ELU receives a negative rate, allowing it to give batch normalization, or mean unit activations close to 0 in Eq. (22).

$$elu(u) = \begin{cases} u \quad if \; u \geq 0 \\ a(e^u - 1) \; if \; u < 0 \end{cases} \qquad (22)$$

The growing network action in the existence of the scaled exponential linear unit (SELU) activation function [26] is analyzed in Eq. (23). Currently, SELU is being suggested during the minor ELU twist. Where λ=1.0507 and α=1.6732, the corresponding equation function is provided.

$$selu(u) = \lambda \begin{cases} u \quad if \; u \geq 0 \\ ae^u - \alpha \; if \; u < 0 \end{cases} \qquad (23)$$

3.4.3 Pooling layer

There is typically a pooling layer after the convolution layer. It decreases the number of parameters and the size of the feature maps, which lowers the cost of the calculation. The assumption of neighboring-pixel estimations exposes the pooling layers to slight modifications. The most used pooling technique is max-pooling. The max-pooling layer therefore often exists through 2 2 as filter size inferred with 2 strides and climbs as high as 4 at most in all convolution layers.

3.4.4 Regularization in deep learning

Making a method that functions well for new entries and training data is a key ML challenge. Several regularization methods for DL have been presented. The dropout is used since it offers stigmatically low-cost services up to an effective method to be regularized. To avoid overfitting, it randomly removes small nodes from the entire connected layer at the training level. As it offers a separate network for training, dropping out is simultaneously considered an ensemble strategy.

3.4.5 Loss function

The key modelling component of a DNN is the choice of the loss function that must be reduced. In general, the categorical cross-entropy function (H) is the best contender and is very useful. It can be explained by two distributions (x and y) over the discrete variable u and is illustrated as in Eq. (24).

$$H(x, y) = -\sum_u x(u) \ln(y(u)) \tag{24}$$

Here, the calculation of the real distribution is indicated by $x(u)$ and $y(u)$.

4. RESULTS AND DISCUSSION

MATLAB R2018b software was used in this investigation and an AMD Quad-Core A10-9620P laptop with a maximum processing speed of 3.4GHz and 4GB of RAM was used to execute it. Training and testing information falls into the 70% training data range and 30% testing data range.

4.1 Image preprocessing evaluation

PSNR (Peak Signal-to-Noise Ratio) and Mean Square Error (MSE):

The parameters that are used to measure the impact of improving image quality in Figure 5. Use Eq. (25) to compare your original image with your final result:

$$MSE = \frac{1}{XY} \sum_{y=0}^{Y=1} \sum_{x=0}^{X=1} (\hat{y}(i,j) - y(i,j))^2 \tag{25}$$

where, $\hat{y}(i,j)$ = matrix of the harvest image and $y(i,j)$ = unique image matrix value.

$$PSNR = 10 \, log_{10} \frac{s^2}{MSE} \tag{26}$$

The higher the PSNR, Eq. (26), the value achieved with S=255 for 8-bit pictures, the better. The MSE was improved as the value got closer to 0.

SSIM (Structural Similarity Index):

For the evaluation of results using quantitative metrics, SSIM is one of the parameters used. When the original image is taken and distortion is added, a vector is created that represents the SSIM value. Use Eq. (27) to calculate the SSIM value.

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{27}$$

where, μ_x = regular value of X, μ_y = regular value of Y, σ_{xy} =correlation X and Y, σ_x^2 = change in values on X, σ_y^2 = alteration in values on Y, C_1 and C_2 are the variables that maintain the squat denominator. Table 1 compares the proposed model with the methods currently in use.

Table 1. Comparative investigation on projected technique with existing methods

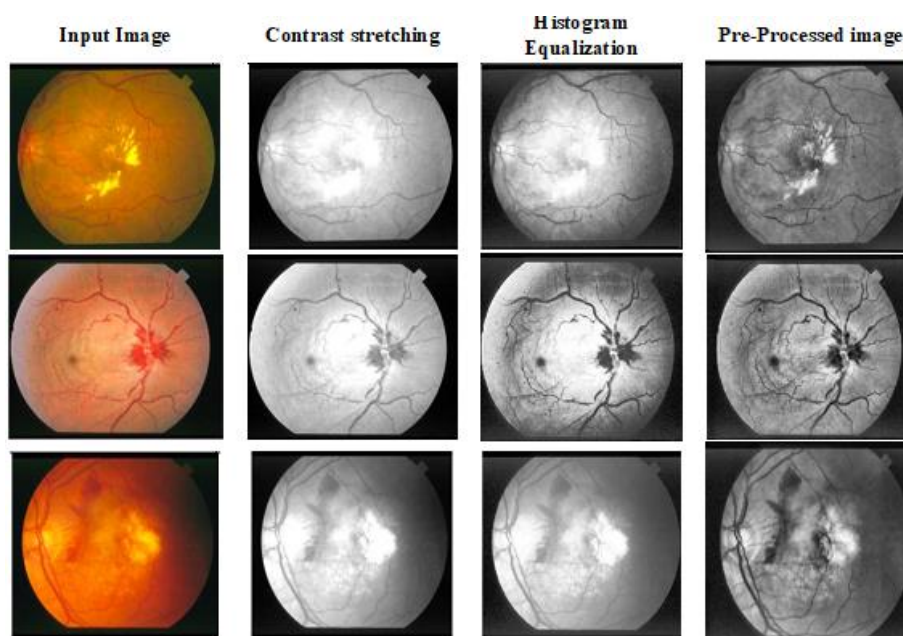| Method	MSE	PNSR	SSIM
Proposed Model	7.33	44.51	0.89
Contrast Stretching+Median Filter [27]	9.14	42.13	0.87
HE+Median Filter [27]	18.68	41.39	0.77
CLAHE+Median Filter [27]	28.42	35.29	0.85
CLAHE+Filtering [28]	-	35.37	0.86
CLAHE+Filtering [29]	85.52	28.88	-
HE [30]	-	25.79	0.56
CLAHE [30]	-	9.48	0.33

Figure 5. Preprocessed and enhanced images after applying the sequence of pre-processing steps

The high PSNR value refers to the performance of the best model; for example, the projected model achieved 44.51% of PNSR, where the existing technique achieved nearly 41% to 42% and CLAHE achieved only 9.48% of PSNR. In the analysis of SSIM, the existing techniques achieved nearly 0.56 to 0.87, but the proposed model achieved 0.89 and this is due to the usage of various steps in preprocessing, as well as segmentation and finally classification is made using CNN model [31-33].

4.2 Classification analysis

PSNR and MSE provide quantitative measures of the difference between the original and enhanced images. In the context of diabetic retinopathy, where subtle features and details in retinal images are critical for accurate diagnosis, having quantitative metrics to assess the fidelity of image enhancement is valuable.

4.2.1 Measure of performance

A performance metric that regularly monitors results and outcomes to generate reliable data on an approach's efficacy is one that is accurate, sensitive, exact, and high in kappa. The kappa index and the general formula for the detection of retinal blood are provided by Eq. (28), Eq. (29), Eq. (30), and Eq. (31).

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \qquad (28)$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \qquad (29)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \qquad (30)$$

$$Kappa\ index = \frac{Accuracy - Accuracy_T}{1 - Accuracy_T} \qquad (31)$$

Here, the terms False Positive (FP), True Negative (TN), False Positive (TP), and True Positive (TP) are all used. The performance of a suggested classifier on two data sets is compared to that of existing techniques in Tables 2 and 3 in terms of metrics.

In the first analysis of the data set, where existing techniques such as RNN, LSTM, and the autoencoder achieved nearly 72% to 92% accuracy, 77% to 88% specificity, 72% to 92% sensitivity, and 80% to 86% kappa index, the projected model achieved 97% accuracy, 98% specificity, 97% sensitivity, and only 87% kappa index. This demonstrates that CNN outperformed other approaches in terms of prediction of DR. The graphical representation of all the strategies that use different metrics for the second dataset is shown in Figure 6.

In the analysis, as shown in Figure 7 and Table 3, of precision, RNN achieved 91%, LSTM achieved 91%, auto-encoder achieved 94%, and CNN model achieved 97%. While tested with sensitivity, RNN achieved 85%, LSTM achieved 89%, auto-encoder achieved 95% and CNN model achieved 97%. The experiments on specificity show that RNN achieved 75%, LSTM achieved 84%, auto-encoder achieved 87%, and CNN model achieved 98%. Compared with all metrics, Kappa index shows less performance, i.e. RNN achieved 79%, LSTM achieved 80%, autoencoder achieved 83% and CNN model achieved 90%.

Table 2. Comparative investigation of proposed classifier with prevailing methods on the first data set

Methodologies	Sensitivity (%)	Specificity (%)	Accuracy (%)	Kappa Index (%)
RNN	72	77	72	86
LSTM	87	83	87	80
Auto-encoder	92	88	92	86
Proposed CNN	97	98	97	87

Table 3. Comparative investigation of the projected classification model with prevailing methods in the second data set

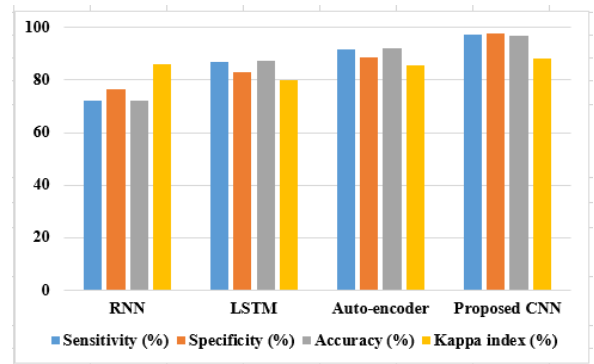| Methodologies	Sensitivity (%)	Accuracy (%)	Specificity (%)	Kappa Index (%)
RNN	85	91	75	79
LSTM	89	91	84	80
Auto-encoder	95	94	87	83
Proposed CNN	97	97	98	90

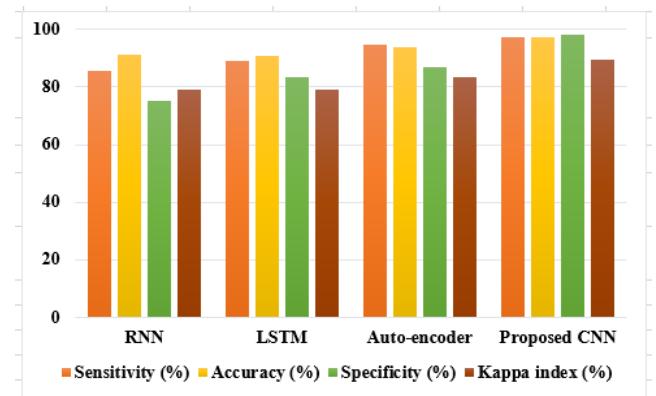**Figure 6.** Representation of projected method using Dataset 1

Figure 7. Representation of the projected model using Dataset 2

5. CONCLUSION

In this investigation, a CNN-based process for the classification of DR was proposed. Patients with DR have blood clots/blobs, exudates, and aberrant blood vessel development in their fundus. This research study employs a series of preprocessing, segmentation, and classification stages. This study applied a series of steps like preprocessing, segmentation, and classification. The main objective of the

study is to eliminate all undesirable noise, size, and color variations in the process, and to enhance the image, the clipping limit, the hyperparameter of CLAHE, is determined using a rat swarm optimization algorithm. The segmentation procedure includes morphological processes, as well as binary conversion and overlay. The classification accuracy is improved by eliminating all photos with small pixel sizes. Finally, the CNN model is used to classify DR pictures. Two datasets, each with its own unique set of metrics, were used to conduct the tests. The results showed that the suggested model outperformed the existing methods in terms of performance. As a result, the proposed model could be improved by incorporating segmentation or feature extraction and selection processes in the future. One limitation of the method "Retinal image enhancement through hyperparameter selection using RSO for CLAHE to Classify Diabetic Retinopathy" lies in its heavy reliance on the hyperparameter optimization process using RSO for CLAHE. Future research could explore the integration of deep learning techniques with the proposed method. CNN have shown remarkable success in various image processing tasks, including medical image analysis. Integrating CNN-based models for both image enhancement and DRs classification could potentially improve accuracy and robustness.

REFERENCES

[1] Ghazal, M., Ali, S.S., Mahmoud, A.H., Shalaby, A.M., El-Baz, A. (2020). Accurate detection of non-proliferative diabetic retinopathy in optical coherence tomography images using convolutional neural networks. IEEE Access, 8: 34387-34397. https://doi.org/10.1109/ACCESS.2020.2974158

[2] Huang, Y.P., Basanta, H., Wang, T.H., Kuo, H.C., Wu, W.C. (2019). A fuzzy approach to determining critical factors of diabetic retinopathy and enhancing data classification accuracy. International Journal of Fuzzy Systems, 21: 1844-1857. https://doi.org/10.1007/s40815-019-00668-0

[3] Ishtiaq, U., Abdul Kareem, S., Abdullah, E.R.M.F., Mujtaba, G., Jahangir, R., Ghafoor, H.Y. (2020). Diabetic retinopathy detection through artificial intelligent techniques: A review and open issues. Multimedia Tools and Applications, 79: 15209-15252. https://doi.org/10.1007/s11042-018-7044-8

[4] Nazir, T., Irtaza, A., Shabbir, Z., Javed, A., Akram, U., Mahmood, M.T. (2019). Diabetic retinopathy detection through novel tetragonal local octa patterns and extreme learning machines. Artificial Intelligence in Medicine, 99: 101695. https://doi.org/10.1016/j.artmed.2019.07.003

[5] Pang, H., Luo, C., Wang, C. (2018). Improvement of the application of diabetic retinopathy detection model. Wireless Personal Communications, 103(1): 611-624. https://doi.org/10.1007/s11277-018-5465-3

[6] Pires, R., Avila, S., Wainer, J., Valle, E., Abramoff, M.D., Rocha, A. (2019). A data-driven approach to referable diabetic retinopathy detection. Artificial Intelligence in Medicine, 96: 93-106. https://doi.org/10.1016/j.artmed.2019.03.009

[7] Rajendran, S., Rajagopal, S.K., Thanarajan, T., Shankar, K., Kumar, S., Alsubaie, N., Ishak, M.K. and Mostafa, S.M. (2023). Automated segmentation of brain tumor MRI Images using Deep Learning. IEEE Access.

[8] Baswaraju, S., Maheswari, V.U., Chennam, K.K., Thirumalraj, A., Kantipudi, M.P., Aluvalu, R. (2023). Future food production prediction using AROA based hybrid deep learning model in agri-sector. Human-Centric Intelligent Systems, 3(4): 521-536. https://doi.org/10.1007/s44230-023-00046-y

[9] Thanarajan, T., Alotaibi, Y., Rajendran, S., Nagappan, K. (2023). Eye-Tracking based autism spectrum disorder diagnosis using chaotic butterfly optimization with deep learning model. Computers, Materials & Continua, 76(2). http://doi.org/10.32604/cmc.2023.039644

[10] Bandello, F., Zarbin, M.A., Lattanzio, R., Zucchiatti, I. (2016). Clinical Strategies in the Management of Diabetic Retinopathy. Springer-Verlag Berlin An, Vol. 10. https://doi.org/10.1007/978-3-642-54503-0

[11] Wang, X., Lu, Y., Wang, Y., Chen, W.B. (2018). Diabetic retinopathy stage classification using convolutional neural networks. In 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, pp. 465-471. https://doi.org/10.1109/IRI.2018.00074

[12] Kumar, S., Adarsh, A., Kumar, B., Singh, A.K. (2020). An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation. Optics & Laser Technology, 121: 105815. https://doi.org/10.1016/j.optlastec.2019.105815.

[13] Das, S., Kharbanda, K., Suchetha, M., Raman, R., Dhas, E. (2021). Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy. Biomedical Signal Processing and Control, 68: 102600. https://doi.org/10.1016/j.bspc.2021.102600

[14] Wang, H., Yuan, G., Zhao, X., Peng, L., Wang, Z., He, Y., Qu, C., Peng, Z. (2020). Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening. Computer Methods and Programs in Biomedicine, 191: 105398. https://doi.org/10.1016/j.cmpb.2020.105398

[15] Leeza, M., Farooq, H. (2019). Detection of severity level of diabetic retinopathy using Bag of features model. IET Computer Vision, 13(5): 523-530. https://doi.org/10.1049/iet-cvi.2018.5263

[16] Hemanth, D.J., Deperlioglu, O., Kose, U. (2020). An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. Neural Computing and Applications, 32(3): 707-721. https://doi.org/10.1007/s00521-018-03974-0

[17] Shankar, K., Sait, A.R.W., Gupta, D., Lakshmanaprabu, S.K., Khanna, A., Pandey, H.M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recognition Letters, 133: 210-216. https://doi.org/10.1016/j.patrec.2020.02.026

[18] Shanthini, A., Manogaran, G., Vadivu, G., Kottilingam, K., Nithyakani, P., Fancy, C. (2021). Threshold segmentation based multi-layer analysis for detecting diabetic retinopathy using convolution neural network. Journal of Ambient Intelligence and Humanized Computing, 1-15. https://doi.org/10.1007/s12652-021-02923-5

[19] Bhimavarapu, U., Battineni, G. (2022). Automatic microaneurysms detection for early diagnosis of diabetic

retinopathy using improved discrete particle swarm optimization. Journal of Personalized Medicine, 12(2): 317. https://doi.org/10.3390/jpm12020317

[20] Islam, M.R., Abdulrazak, L.F., Nahiduzzaman, M., Goni, M.O.F., Anower, M.S., Ahsan, M., Haider, J., Kowalski, M. (2022). Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images. Computers in Biology and Medicine, 146: 105602. https://doi.org/10.1016/j.compbiomed.2022.105602

[21] Patry G, Gauthier G, Bruno LAY, Roger J, Elie D, Foltete M, Donjon A, Maffre H ADCIS, 2018

[22] Riya, K.S., Surendran, R., Tavera Romero, C.A., Sendil, M.S. (2023). Encryption with user authentication model for internet of medical things environment. Intelligent Automation & Soft Computing, 35(1). http://dx.doi.org/10.32604/iasc.2023.027779

[23] Dutta, S., Manideep, B.C., Basha, S.M., Caytiles, R.D., Iyengar, N.C.S.N. (2018). Classification of diabetic retinopathy images by using deep learning models. International Journal of Grid and Distributed Computing, 11(1): 89-106. http://doi.org/10.14257/ijgdc.2018.11.1.09

[24] Dhiman, G., Garg, M., Nagar, A., Kumar, V., Dehghani, M. (2021). A novel algorithm for global optimization: rat swarm optimizer. Journal of Ambient Intelligence and Humanized Computing, 12(8): 8457-8482. https://doi.org/10.1007/s12652-020-02580-0

[25] Houssein, E.H., Hassaballah, M., Ibrahim, I.E., AbdElminaam, D.S., Wazery, Y.M. (2022). An automatic arrhythmia classification model based on improved marine predators algorithm and convolutions neural networks. Expert Systems with Applications, 187: 115936. https://doi.org/10.1016/j.eswa.2021.115936

[26] Manzo, M., Pellino, S. (2021). FastGCN+ARSRGemb: A novel framework for object recognition. Journal of Electronic Imaging, 30(3): 033011. https://doi.org/10.1117/1.JEI.30.3.033011

[27] Ningsih, D.R. (2020). Improving retinal image quality using the contrast stretching, histogram equalization, and CLAHE methods with median filters. International Journal of Image, Graphics and Signal Processing, 10(2): 30-41. http://doi.org/10.5815/ijigsp.2020.02.04

[28] Sahu, S., Singh, A.K., Ghrera, S.P., Elhoseny, M. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Optics & Laser Technology, 110: 87-98.

https://doi.org/10.1016/j.optlastec.2018.06.061

[29] Zulfahmi, R., Noviyanti, D.S., Utami, G.R., Harison, A.N., Agung, P.S. (2019). Improved image quality retinal fundus with contrast limited adaptive histogram equalization and filter variation. In 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, IEEE, pp. 49-54. https://doi.org/10.1109/ICIMCIS48181.2019.8985198

[30] Gupta, B., Tiwari, M. (2019). Color retinal image enhancement using luminosity and quantile based contrast enhancement. Multidimensional Systems and Signal Processing, 30(4): 1829-1837. https://doi.org/10.1007/s11045-019-00630-1

[31] Selvanarayanan, R., Rajendran, S., Alotaibi, Y. (2024). Early detection of colletotrichum kahawae disease in coffee cherry based on computer vision techniques. CMES-Computer Modeling in Engineering & Sciences, 139(1). http://doi.org/10.32604/cmes.2023.044084

[32] Selvanarayanan, R., Rajendran, S., Algburi, S., Ibrahim Khalaf, O. and Hamam, H., (2024). Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system. Scientific Reports, 14(1): 6269.

[33] Alharbi, M., Rajagopal, S.K., Rajendran, S., Alshahrani, M. (2023). Plant disease classification based on ConvLSTM U-Net with fully connected convolutional layers. Traitement du Signal, 40(1): 157-166. https://doi.org/10.18280/ts.400114

NOMENCLATURE

CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
DR	Diabetic Retinopathy
FOV	Field of View
IDRiD	Indian Diabetic Retinopathy Image Dataset
MESSID	Methods to Evaluate Segmentation and Indexing
OR	Techniques in Field of Retinal Ophthalmology
MA	Microaneurysm
NPDR	Non-Proliferative Diabetic Retinopathy
PDR	Proliferative Diabetic Retinopathy
RSO	Rat Swarm Optimization
AOSLO	Adaptive optics scanning laser ophthalmoscopy