

Original Article

# An Optimization Method for Tuning Hyper-Parameters of SGAN with Ensemble Classification Regression Model

Valarmathi Srinivasan<sup>1</sup>, Vijayabhanu Rajagopal<sup>2</sup>

<sup>1,2</sup>Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education for Women, Tamilnadu, India

<sup>1</sup>Corresponding Author : valarmathis313@gmail.com

Received: 18 January 2023

Revised: 15 March 2023

Accepted: 03 April 2023

Published: 27 April 2023

**Abstract** - Diabetic Retinopathy (DR) is a potential condition of Diabetes Mellitus (DM) that causes lesions on the retina, reducing vision and perhaps leading to disability if not properly treated. A Special Generative Adversarial Network with Ensemble Classification Regression (SGAN-ECR) model was introduced for the effective categorisation of various DR grades with the generation of high-contrast and low-saturated RF images. However, the Multi-Scale Attention Residual Network with Gradient Boosting (MSA-ResNet-GB) model used in SGAN-ECR requires manually assigning values for many hyper-parameters. An inappropriate value of the hyper-parameter leads to an increased error rate. This paper presents a method using the Enhanced Mine Blast Algorithm (EMBA) for selecting optimal hyper-parameters of MSA-ResNet-GB, such as the number of convolution layers, number of filters, filter size, number of Fully Connected (FC) layers and the hidden units in the FC layer to improve the architecture. Essential principles of EMBA are derived from the mine bomb explosion in real-time applications. The initial population of shrapnel fragments represents the initial hyper-parameter, and the computation of their subsequent locations represents the search for the best hyper-parameter. SGAN-ECR with parameter-optimised MSA-ResNet is named SGAN-OECR. In SGAN-OECR, optimised MSA-ResNet is trained by training images, and then the trained model is used to recognise the abnormalities in test images. MLP classifiers in the last FC layers of MSA-ResNet classify the severity of DR lesions. Finally, the experimental results of SGAN-OECR on Kaggle-APTOS and IDRiD datasets have 99.03% and 98.63% of accuracy, respectively, which is higher than the existing DR classification models.

**Keywords** - Diabetic retinopathy, Generative adversarial network, Hyper-parameters, Mine blast algorithm, Optimisation, Retinal fundus images.

## 1. Introduction

Diabetes is one of the most prominent causes of DR, which greatly impairs eyesight. It fluctuates in severity and is medicated when detected early [1]. The retina transforms light into an electrical signal that can be altered to generate an image. The retina is surrounded by a network of blood arteries supplying oxygen and nutrients. Diabetes alters blood vessels, which reduces blood supply to the retina. This impairs the retina's health and reduces visual acuity. The most chronic kind of DR is non-proliferative retinopathy. At this stage, diabetes has minimal effects on eyesight, despite altering blood vessels. DR lesions include arteries that expand considerably (microaneurysms), leak fluid and proteins (exudates) or bleed blood (retinal haemorrhages) [2].

It is expensive to detect DR manually by ophthalmologists or professional graders. Various computeraided DR detection and diagnosis strategies have

been developed in the past year. Computer-Aided Detection (CADE) methods identify and discriminate lesions at the pixel level, while Computer-Aided Diagnosis (CADx) models distinguish DR at the image level [3,5]. Evaluators can use the CADE and CADx systems to help them take patients to an ophthalmologist. Machine Learning (ML) and Deep Learning (DL) algorithms have been used to execute alternative strategies for CADx systems. Scale-invariant features extracted from DR lesion images are used to identify visual interest points. Probabilistic latent semantic analysis reduces feature dimensions. Support vector classifiers identify images as normal or pathological using visual terms [6].

ML does not dynamically produce features from the dataset that uses only the given features. The given features are either inaccurate representations of tumor forms in RF images or are not fitted to them [7]. Recently, DL has been used to estimate DR from RF pictures, and the results are



promising [8,9] and address the difficulties mentioned above. Convolutional Neural Network (CNN) is a popular DL model comprising millions of base learners and requires a high number of RF scans to train in the case of DR, resulting in generalisation error. Some advanced and pre-learned CNN architectures like VGG [11], ResNet [11] and Dual Path Network (DPN) [12] are used for DR lesion detection. These models efficiently transform low-level features into high-level features. The pre-learned model produced the best results for even fewer training samples. However, these models failed to properly categorise some pixels at the edges of tumours and healthy regions.

ResNetGB [13] is a pre-trained CNN structure constructed to train the Regions-of-Interest (ROI) in tumours from pixel-level tumor labelling. The layout was designed to train the tumor and normal neighbourhood patterns. A GB-based classifier was also included in determining the DR values of RF images. Multiple layers of features are required to explain the DR severity stage, despite the fact that the highest range of features are retrieved to identify and label retinal lesions as DR. To solve this, MSA-ResNetGB [15] model was developed to extract a mix of mid and high-level features to reinforce the interpretation. The Multi-Scale Feature Pyramid (MSFP) was also constructed to define the RF pattern in various positions GB classifier was used to classify DR grades. Although this model performs well for RF images, the accuracy is ineffective for low-contrast and saturated RF images. A Special Generative Adversarial Network with Ensemble Classification Regression (SGAN-ECR) model was developed to generate high-contrast and low-saturated RF images from low-contrast and saturated RF images for efficient training. In SGAN, pixel-2-pixel GAN and UNet++ are used in the generator and patch-GAN is used in the discriminator. MSA-ResNet was used to retrieve features. Multi-Layer Perceptron (MLP) algorithm instead of GB is used to categorise the different phases of DR severity. However, a classifier's accuracy mainly depends on assigning values for hyper-parameters used in MSA-ResNet.

Hence, an EMBA optimisation method is developed in this article to optimise the hyper-parameters of the MSA-ResNet structure in the SGAN-ECR model. The optimal hyper-parameters, such as the number of convolution layers, No. of filters, filter size, No. of FC layers, and the hidden units in the FC layer are selected by using the EMBA scheme. The initial population of shrapnel pieces of EMBA is analogous to the initial hyper-parameters. Their updated location computation is analogous to searching for the best hyperparameters. The updated locations' fitness values (classification error) are improved iteratively until all individuals have identical fitness values. Since SGAN-ECR is reconstituted with parameter-optimised MSA-ResNet is

referred to as SGAN-OECR. The presented model improves the classification of DR severity from RF images.

This manuscript is arranged as Section II reviews the previous research associated with DR severity. Section III explains the SCAN-OECR algorithm for categorising the DR severity from RF images by optimising the hyper-parameter in MSA-ResNet using EMBA, and Section IV displays its efficacy. Section V summarises the whole study and suggests further development.

## 2. Literature Survey

A DL interpretable classifier was presented [16] to classify the DR grades. This model determined categorisation results by applying a score to every hidden and input data. Such scores represent the importance of each pixel to the objective categorisation. These results were obtained by utilising a new framework for the pixel-wise transmission of results across all neurons. However, efficiency was diminished due to an absence of DR images with subjectively designated lesions.

A new hybrid model was introduced [17] based on image processing and CNN layer to identify and classify DR from retinal fundus images. At first, pre-processing was performed using histogram equalisation and the CLAHE to enhance the input images for better classification. Then, the CNN was applied to classify those images to detect the DR. But, it needs additional databases and new pre-processing methods to increase the efficiency.

The decision tree-based ensemble learning (DTEL) method was introduced as a novel technique [18] for DR identification using gray-level intensity and texture data collected from fundus pictures. First, images containing noise were filtered out using pre-processing. The histogram and gray-level co-incident matrix of every image was computed to get the statistical features. In addition, the genetic technique was used to identify the most relevant features, which were then taught by the XGBoost algorithm to estimate the intensity of DR. However, the model parameters must be fine-tuned to improve efficiency.

A new hybrid method was presented [20] for prior DR identification and categorisation. First, the collected images were pre-analysed using the adaptive histogram equalisation, contrast stretching and median filtering schemes. After, enhanced images were given to the bright and red lesion recognition schemes to extract the features on the recognised areas. Moreover, those features were classified by the SVM, KNN and binary trees along with the combined voting scheme. On the other hand, the pre-processing stage was time-consuming because of engaging all image-processing phases.

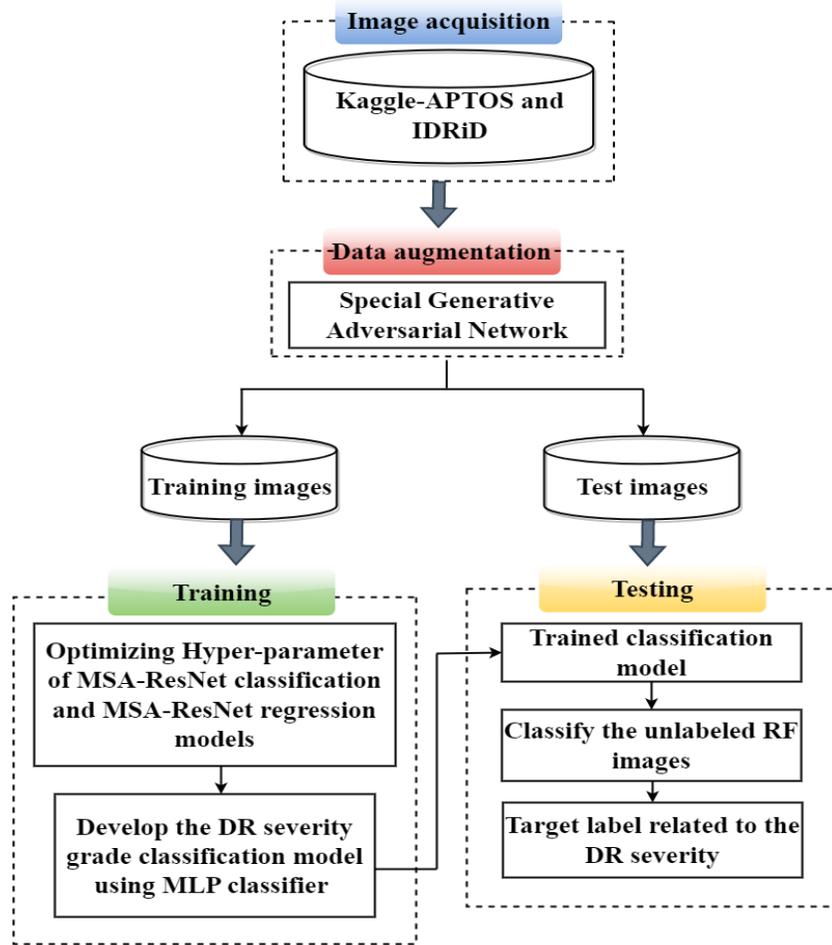


Fig. 1 Flow diagram of the SGAN-OECR model for DR disease classification

A VGG-NIN method was developed [21] to classify the different DR stages with the minimum feasible trainable variables to accelerate learning and model convergence. This model integrated the VGG16, spatial pyramid pooling layer, and network-in-network to produce a very indiscriminate, persistent structure for deep identification of DR images. However, the accuracy was less since the pre-processing steps were not effective.

An automated DR grading model was developed [22] depending on identifying multiple retinal lesions. Initially, the retinal images were pre-processed, and the grey-level run-length matrix average was determined to distinguish normal and DR images. Then, the U-Net was used to extract the exudates, microaneurysms, haemorrhages and blood vessels. Also, different characteristics were retrieved and classified by the multi-mode SVM to classify the DR grades. But, it needs to consider the different disease features concurrently to increase the accuracy.

Deep Transfer Learning (DTL) was introduced [23] based on the different pre-trained CNN models such as "AlexNet, GoogleNet, InceptionV4, Inception ResNetV2 and

ResNext50" to detect Retinopathy automatically. First, the DR images were acquired and labelled with a suitable diagnosis method. Then, the different CNN models were trained to classify the severity of DR images. However, a more in-depth analysis was required, which needed a massive amount of patient information.

A Model was developed [25] for recovering retinal images using a cycle-consistent generative adversarial network (CycleGAN) with a convolution block attention component. Using a modified UNet, the retinal vessels in the regenerated retinal images were segmented. This model has two generators and two discriminators. Generators convert low-quality images from one classification to another and vice versa. Discriminators classify real and manufactured images. In order to get segmented images, the model for retinal vascular segmentation incorporates downsampling, bottlenecking, and up-sampling layers. CBAM was utilised to enhance the feature extraction of these models. This model demands a considerable amount of data training time.

A DL protocol Funswim was created [26] to determine the rate of DR and the risk of retinal edema based on FIR. The

swim transformer works remarkably well in this technique for categorising medical images. The ImageNet-based TL model was applied as a training model to collect sufficient low-level features for learning. Therefore, when the model was combined with high-dimensional features, the systems could be analysed more accurately (particularly potential disease classification bases) and improve classification efficiency. However, this model generated a slower convergence rate.

A Deep CNN (DCNN) model for autonomous identification and categorising DR from color FIR was described [27]. Each image was standardised using Min-Max standardisation to minimise the CNN from training the image's underlying background noise. The CNN was then utilised to categorise and identify the DR. Finally, dynamically, the quadratic kappa metric was employed to assess the accuracy of the predictions. However, sample sizes for all cohorts were rather small.

A Hybrid DL (HDL) framework was devised [28] to recognise and categorise DR in the FIR of the eye. CNN models were subjected to TL to obtain features from ResNet-18 and GoogleNet models, which were then connected to generate a composite feature vector. Using these attributes as input, multiple classifiers performed binary and multiclass categorisation of DR images from the acquired dataset. Integrating GoogleNet and ResNet-18 features increases the aggregate multiclass recognition efficiency of the system. However, this system has a slower processing speed.

### 3. Proposed Methodology

In this part, the SGAN-OECR model is described briefly. Figure 1 depicts the flowchart of the provided hyper-parameter maximising system for classifying the various degrees of DR severity from RF images. Primarily, different RF image databases are acquired and then the SGAN-ECR is applied to get the number of high-quality RF images. Such images are further trained by the classification and regression models using MSA-ResNet structure [15] with MLP classifier. In addition, the generated model is deployed to the test samples to categorise the various severity levels of DR. The following sections describe these different processes.

#### 3.1. MSA-ResNet Structure Optimization

This framework intends to propose an efficient method for autonomously optimising the hyper-parameters and developing an MSA-ResNet for DR disease classification. The efficiency of the MSA-ResNet architecture is extremely dependent on the hyper-parameter values. Since optimising all MSA-ResNet hyper-parameters would require considerable time. The number of convolution layers, the number of filters, the filter size, the number of FC layers, and the hidden units in the FC layer are thereby optimised thoroughly in this suggested framework.

**Table 1. Configuration of hyper-parameters for MSA-ResNet structure**

Model	Hyper-parameter	Value
MSA-ResNet	Conv_1 layer	7*7
	Conv_2 layer	3*3
	Batch Size	4
	Epochs	100
	Learning Rate	$10^{-4}$
	Optimizer	Adam Optimizer
	Activation Function	Sigmoid
	Pooling	Adaptive and Global Mean Pooling
	Residual Blocks	3, 8, 12, and 3

The optimal MSA-ResNet structure is expressed as the following list of hyper-parameters:

$$H = H_{conv}, H_{fc} \quad (1)$$

From above Equation (1),  $H_{conv}$  stands for the parameters of the convolutional layer and  $H_{fc}$  stands for the parameters of the FC layer. The convolutional layer's parameters are defined as follows:

$$H_{conv} = \{c_0, \dots, c_{N-1}\} \quad (2)$$

Where  $c_N$  represents the number of convolutional layers in Equation (2). Every Convolutional layer in the  $a^{th}$  layer consists of two tuples,

$$c_a = (d_{count}, d_{size}) \quad (3)$$

Where, the  $d_{count}$  and  $d_{size}$  are the No. of filters and size of the filters of the  $a^{th}$  layer, respectively, in equation (3). Equation (4) formulates the set of the fully-connected layers,

$$H_{FC} = t_0, t_{k-1} \quad (4)$$

Where,  $t$  represents the hidden division number in the FC layer of the  $a^{th}$  layer, and  $k$  is the No. of FC layers.

The possible set of MSA-ResNet configurations is denoted by  $\mathcal{Q}$ . The objective is to identify a configuration  $q \in \mathcal{Q}$  which provides a minimum classification error rate. In this proposed model, the Classification error rate ( $Cer$ ) is considered the objective function to enhance the classification accuracy. EMBA selects the possible hyper-parameter values in this research work. The selection of optimal parameters reduces the utilisation of computer resources. Table 1 depicts the Hyper-parameter and its range of MSA-ResNet.

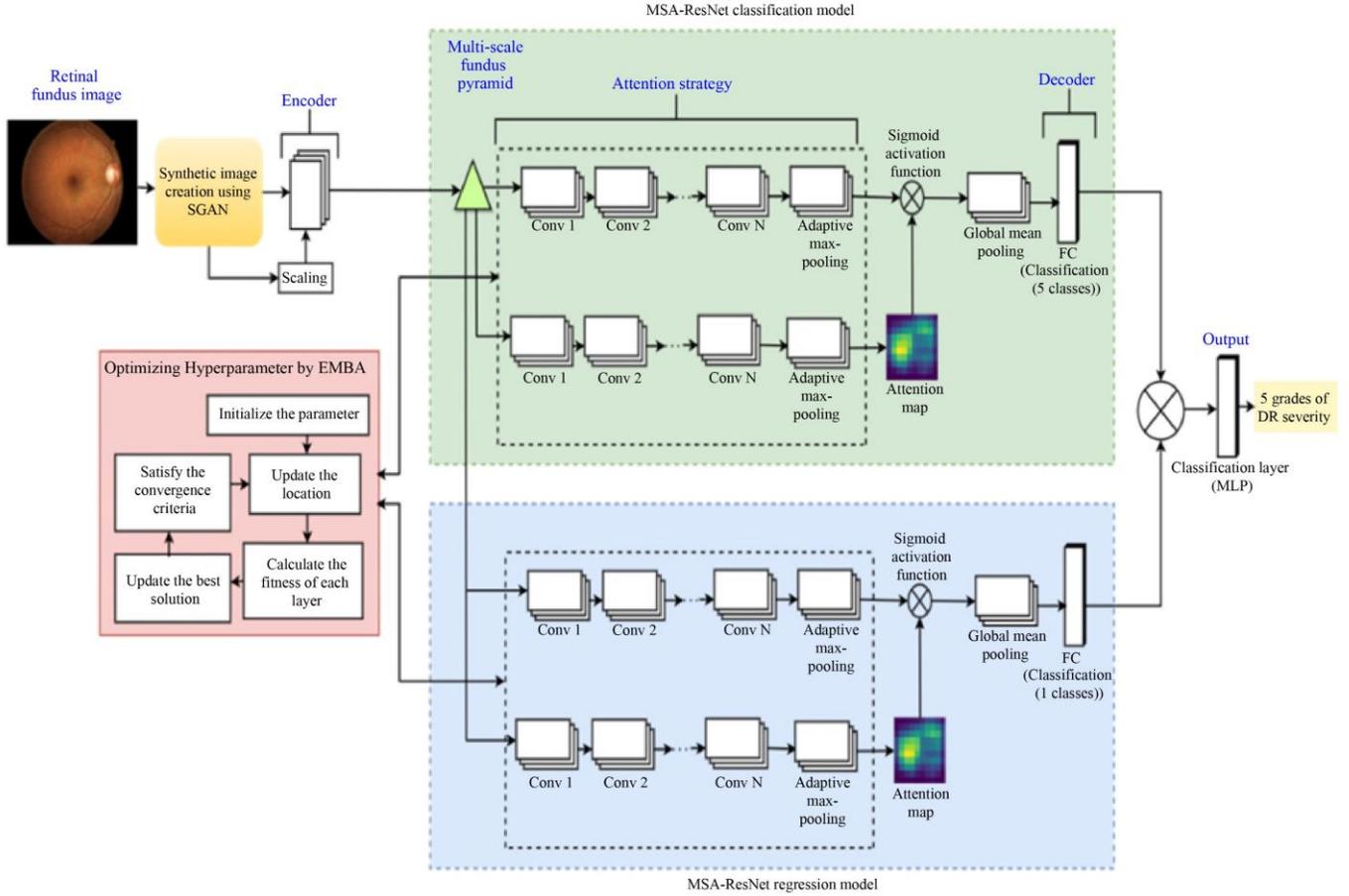


Fig. 2 Schematic representation of SGAN-OECR model

### 3.2. EMBA for Optimising Hyper-Parameters in MSA-ResNet Structure

The structure of the SGAN-OECR model is provided in Figure 2. This is the optimised version of SGAN-ECR. SGAN-OECR comprises of SGAN and ECR. SGAN has GAN with UNet++ as a generator and patch-GAN as a discriminator for generating quality images to get a highly relevant training model from the classifier. The ECR part has MSA-ResNet regression and MSA-ResNet classification structures for feature extraction, training and classifying for DR severity grades. The EMBA algorithm in the proposed work optimises the MSA-ResNet model's hyperparameters. An appropriate selection of parameters provided high accuracy in less number of iterations.

MBA is one of the most recently developed metaheuristic algorithms, and EMBA is an upgraded version of MBA. This algorithm has given superior outcomes when compared to other metaheuristic optimisation techniques. The explosion of landmines influenced this MBA method. In summary, the MBA claims that when a landmine explodes, the resulting shrapnel disperses and collides with other landmines, destroying them. In order to achieve the optimal solution, it is essential to identify the initial blasting point at which

sufficient collisions are generated to remove the region of mines completely. Consequently, when an explosion occurs at an initial location( $X_0$ ). Shrapnel fragments are produced and disseminated over the (search) zone for the first time. As a result, more explosions occur when these shrapnel particles interface with surrounding landmines. The position of every explosion differs from the position of the preceding explosion, and the number of casualties is generated and represented by  $f(X)$  for each point ( $X$ ). This method is a population-based technique, which necessitates the generation of a starting population. This population (which is the ideal alternative at the time) is generated by the initial blast (first shot), and its size is determined by the number of shrapnel fragments generated by the original shot. The two parts of this MBA procedure are exploration (global search) and exploitation. The proliferation of shrapnel prompted the examination stage to numerous additional regions, whereas the exploitation stage comprises exploiting the existent resources in these designated sites. Figure 3 displays a diagrammatic description of the MBA in terms of exploration (color lines) and exploitation (black color lines) processes, where  $x_0, \dots, x_5$  represent the beginning positions and the subsequent positions.

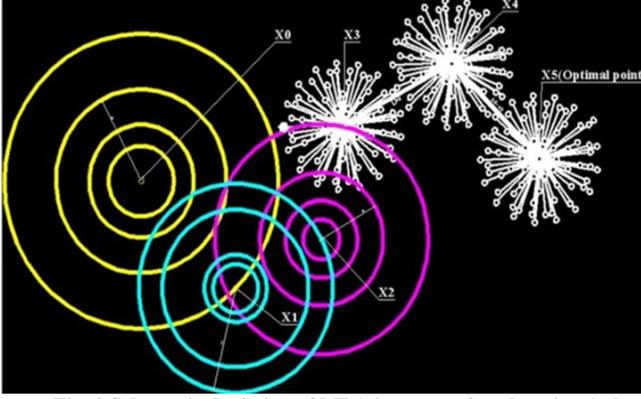


Fig. 3 Schematic depiction of MBA in terms of exploration (color lines) and exploitation (white color lines) operations

A population generation is here represented as the analogy of generating the set of MSA-ResNet hyperparameters configurations  $Q$ , and the solutions or fitness of each population is represented as,

$$f(Q) = f(X) = \frac{1}{1+|Cer|} \quad (5)$$

In Equation (5),  $f(Q)$  represent the fitness or objective value for the configuration set  $Q$ .  $f(X)$  represents the fitness of exploitation at shot point  $X$ . Here  $X$  is the analogy for the configuration set  $Q$ . So, in the following sections,  $X$  represent the configuration set for hyperparameters of MSA-ResNet.  $Cer$  is the classification error of MSA-ResNet for a selected configuration parameter.

### 3.2.1. Steps Involved in MBA

The initial shot point (hyperparameter set  $Q_s$  for MSA-ResNet) determines the response space to be searched, and this initial shot point is chosen arbitrarily. This EMBA uses the lower limit parameters of a given problem as though the location of the original shot point is irrelevant and performs a fresh first shot. EMBA employs a changing randomly completed value using Equation (6),

$$X_0^{current} = lb + rand * (ub - lb) \quad (6)$$

Suppose,  $X_0^{current}$  is the point of the initial explosion point,  $lb$  and  $ub$  depicts lower bound upper bound search space, respectively. This means the minimum and maximum range of values for the hyperparameter set  $Q_s$  for MSA-ResNet. The next explosion location (the updated configuration sets for MSA-ResNet) will occur later, which will be represented as,

$$X = X_T, T = 1, 2, 3, \dots, n_d \quad (7)$$

Where,  $n_d$  is a search space for positions  $X$  in Equation (7). Assume further locations of the detonated landmine, as identified by the fragments (shrapnel), using Equation (8),

$$X_{n+1}^b = X_{z(n+1)}^b + \exp\left(-\sqrt{\frac{T_{n+1}^b}{d_{n+1}^b}}\right) * X_n^b \quad (8)$$

$$n = 0, 1, 2, \dots, d$$

$X_{z(n+1)}^b$  represents the location of the exploding mine bomb, whereas  $T_{(n+1)}^b$  and  $d_{(n+1)}^b$  are the orientation and range of the shrapnel pieces accomplished in every iteration. The subscript  $b$  indicates the total amount of initial shot points.

Determine the exploding landmine location by Equation (9) as flows,

$$X_{z(n+1)}^b = d_{n+1}^b \times r \times \cos(\theta) \quad (9)$$

Where,  $r$  is the random integer from 0 to 1 and is the shrapnel angle. This angle  $\theta$  corresponds to  $360/N_b$ .

The distance and the directions of shrapnel pieces are computed by equations (10) and (11), and the position of the exploded landmine is calculated by utilising equation (9).

$$d_{n+1}^b = \sqrt{(X_{n+1}^b - X_n^b)^2 + (F_{n+1}^b - F_n^b)^2} \quad (10)$$

$$T_{n+1}^b = \frac{F_{n+1}^b - F_n^b}{X_{n+1}^b - X_n^b} \quad (11)$$

$F$  is the value of the fitness function at the location of  $X$ . In the exploitation phase, the EMBA uses a fitness value of the current solution. In order to attain a balance between these two random operations, a fitness function ( $F$ ) from Eq. (5) is used in Eqs. (10) and (11). The fitness function represents the classification error of MSA-ResNet for the parameters selected at location  $X$ .

Two techniques are employed to enhance the exploration and exploitation of EMBA.

### The Exploration Process:

The user specifies an initial distance, which is utilised to search for the optimal solution within a range determined by the ratio of the original length and an arbitrarily determined number. In addition, exploration is incorporated to perform exploration of the layout space at lesser and greater distances. This factor is utilised in the early iterations of the method, compared to an iteration number. If it is more than the iterative number, the resulting space is explored as follows:

$$d_{n+1}^b = X_n^b * r^2 \quad (12)$$

$$X_{z(n+1)}^b = d_{n+1}^b * \cos(\theta) \quad (13)$$

### The Exploitation Process:

This procedure is determined by the exploration factor ( $\gamma$ ) if it is lower than the number of iterations ( $b$ ). By minimising the constant value, the distance traversed by fragments in an explosion is decreased in this instance. Candidates will obtain the fixed value. This elimination in the distance is calculated by Eq. (13) as follows:

$$X_n^b = \frac{\gamma_{n-1}^b}{\exp\left(\frac{b}{e}\right)} \quad (14)$$

### Convergence Criteria:

In the majority of metaheuristic algorithms, the optimal result is obtained when the terminating criteria are considered to be the greatest iterative values, the query distance, or a smaller value set as the permitted deviation among the previous two outcomes. This is the primary distinction between EMBA and MBA. The explosion positions are updated until the below convergence criteria are satisfied.

The difference in the length among the current exploded point is  $X_{z(n+1)}^b$  and the greatest solution will be  $X_{best}$ . The enhancement is considered as follows

$$X_{n+1}^b = X_{z(n+1)}^b + \exp\left(-\sqrt{\frac{1}{ED}}\right) \times r \otimes \{X_{best} - X_{z(n+1)}^b\},$$

$$j = 1, 2, \dots, n \quad (15)$$

$$ED = \left[ \sum_{i=1}^T (X_{best} - X_{z(n+1)}^b)^2 \right]^{1/2} \quad (16)$$

Where  $ED$  indicates the Euclidean distance between the optimal solution location  $X_{best}$  and the current explosion point  $X_{z(n+1)}^b$  in  $T$  dimensions. The current method did not utilise preceding optimal location information, hence accelerating the convergence of the algorithm. The general pseudo code of SGAN-OECR is given below.

### 3.3. Pseudo Code for SGAN-OECR Model

**Input** : DR lesions and healthy images for 5 different DR stages.

//Images from Kaggle-APTOS and IDRiD for 5 different DR grades; Get the RF images Perform SGAN model to create the synthetic RF images and add them to the real RF image databases; Divide the newly obtained RF image databases into the training set and test set;

**Output** : Different grades of DR severities

#### Begin

**Step 1** : Choose the preliminary parameters:  $N_d$ ,  $\gamma$ ,  $\varepsilon$ ,  $ub$ ,  $lb$  and maximum number of iterations  $p_{max}$  for EMBA.

**Step 2** : *while* ( $p < P_{max}$ )

**Step 3** :  $P \leftarrow p + 1$ ;

**Step 4** : Initialize the initial shot points using Equation (6) //this may cause the algorithm to explore alternative regions in the resulting space. For an analogy, this point has hyperparameter configuration set  $Q$  for MSA-ResNet, i.e. No. of convolution layers, No. of filters, filter size, No. of FC layers, hidden units in the FC layer.

**Step 5** : Generate other shrapnel pieces using Equation (7), Equation (8) and Equation (9) and update their locations.

**Step 6** : Compute the distance and directions of shrapnel pieces direction by equations (10) and (11).

**Step 7** : Calculate the fitness of all location and direction updated shrapnel pieces using MSA-ResNet classification and regression, saving the highest fitness as the best temporal shrapnel piece.

**Step 8** : Improve the shrapnel pieces' position by exploration-exploitation from Equation (12) to Equation (14)

**Step 9** : Examine the convergence criteria. If the stopping requirement is met, the algorithm will be terminated. Otherwise, proceed to step 5.

**Step 10** : If true, save the improved shrapnel piece as the best temporal response

**Step 11** : If not, replace the shrapnel's location with the optimal temporal response

**Step 12** : Examine the convergence criteria. If the stopping rule is satisfied, the algorithm will be terminated. Else, proceed to step 5.

**Step 13** : End while

**Step 14** : Select highest fitness shrapnel pieces and obtain the trained model from the MSA-ResNet classification model;

**Step 15** : Classify the test images into different grades of DR severity using the trained model;

**End**

Table 2 depicts the MBA parameters and their value for performing MSA-ResNet and GAN.

Table 2. MBA parameter settings

Parameter	Value
Upper bound	1
Lower bound	0
Maximum distance	2
Number of iterations	200
Population size	20
KNN Classifier	5

## 4. Experimental Results

The SGAN-OECR model is assessed using four different metrics like Accuracy, Precision, Recall, and F1-Score using benchmark datasets, Kaggle-APTOS and IDRiD.

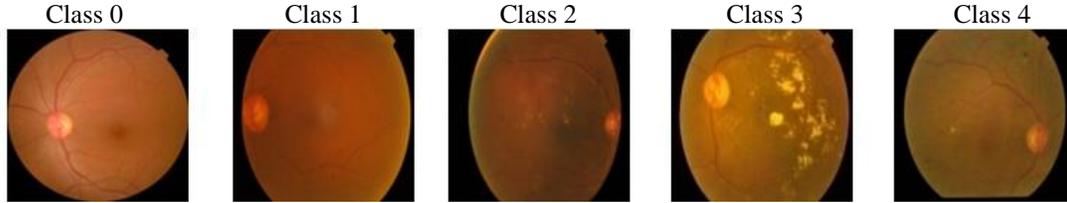


Fig. 4 Sample RF images taken from the Kaggle-APTOS corpus

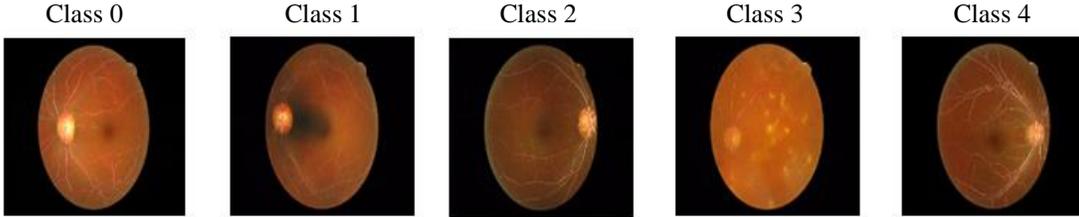


Fig. 5 Augmented RF images taken from the Kaggle-APTOS corpus

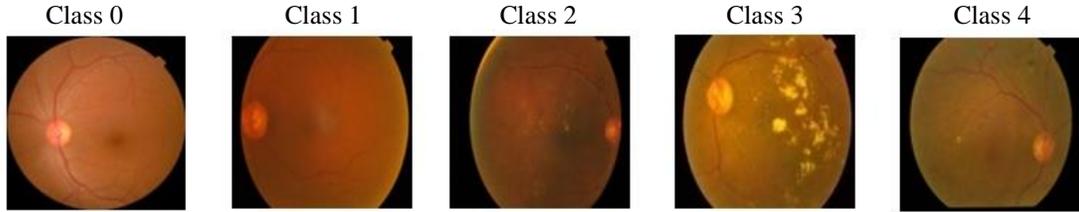


Fig. 6 Sample RF images taken from the IDRiD

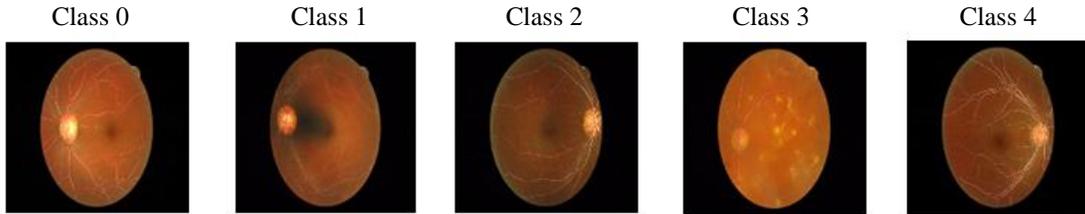


Fig. 7 Augmented RF images taken from the IDRiD

#### 4.1. Dataset Description

The proposed model was assessed using benchmark datasets, Kaggle-APTOS and IDRiD. A number of samples that were gathered, the samples that were annotated, the DR grades, and the ground truths are all detailed in this subsection. From both datasets, sample RF pictures are displayed.

##### 4.1.1. APTOS 2019 Blindness Detection Database

This study used RF images from the "Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 Blindness Detection dataset" [29]. This Kaggle image collection has 3662 images obtained from various rural Indians. "Aravind Eye Hospital" in India gathered and organised the data. However, a group of medical specialists used the International Clinical Diabetic Retinopathy Disease Severity Scale to examine and classify the obtained samples.

APTOS dataset images are divided into five classes on a scale of 0 to 4: "no DR, mild DR, moderate DR, severe DR,

and proliferative DR". The initial class includes healthy RF samples that do not include DR. The preceding classes have more injured retinas than the previous ones. The final categorisation was proliferative DR, which includes instances with vitreous or pre-retinal HEs. Figure 4 illustrates the RF samples from every Kaggle-APTOS class, whereas Figure 5 depicts the enhanced RF samples from each Kaggle-APTOS class.

##### 4.1.2. Indian Diabetic Retinopathy Image Database (IDRiD)

Images from the IEEE ISBI-2018 are used in the "Indian Diabetic Retinopathy Image Database (IDRiD)" [30]. It includes 516 images with medical indications for DR and DME, containing 413 training images and 103 test images under the heading "Disease grading" as appropriate. IDRiD was the first dataset to represent the Indian samples. Figure 6 shows the RF images from every category in the IDRiD, whereas Figure 7 shows the enhanced RF samples from every class in the IDRiD.

**Table 3. Performance analysis of the proposed model and the cutting-edge models for DR classification on "APTOS 2019 blindness detection dataset"**

Ref.	Year	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[18]	2021	DT-EL	86.81	87.31	87.74	87.51
[13]	2021	ResNetGB	88.40	88.72	88.40	88.47
[21]	2021	VGG-NIN	87.65	92.86	92.86	92.86
[22]	2021	U-Net	87.37	91.76	93.98	92.86
[26]	2022	Funswim	88.89	92.59	94.94	93.75
[27]	2022	DCNN	89.47	94.12	94.12	94.12
[15]	2022	MSA-ResNetGB model	94.40	94.52	94.40	94.42
-	-	SGAN-ECR	97.81	95.95	97.98	96.90
Proposed Model	2023	SGAN-OECR (Conventional MBA)	97.92	96.97	98.12	97.85
		SGAN-OECR (Enhanced MBA)	98.63	98.83	98.28	98.55

**Table 4. Performance analysis of the proposed model and the existing models for DR classification on the "IDRiD Database"**

Ref.	Year	Classification method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
[17]	2020	CNN model	86.02	89.67	90.24	91.93
[20]	2021	Mixed Model	87.50	91.76	93.38	92.86
[13]	2021	ResNetGB	88.89	93.51	93.51	93.51
[23]	2022	DTL model	89.01	93.83	93.83	93.83
[25]	2022	CBAM-UNet	89.47	94.94	92.59	93.75
[28]	2022	HDL Model	90.11	93.83	95.00	94.41
[15]	2022	MSA-ResNetGB model	94.17	91.48	91.57	91.45
-	-	SGAN-ECR	96.12	96.31	93.73	94.77
Proposed Model	2023	SGAN-OECR (Conventional MBA)	98.61	98.26	95.92	96.89
		SGAN-OECR (Improved MBA)	99.03	99.39	96.00	97.47

The prevalence of DR and DME is reported at the pixel level for every instance in the IDRiD collection. Based on the intensity scale, the DR grade was categorised into five groups from 0 to 4, equivalent to the Kaggle-APTOS dataset.

#### 4.2. Performance Metrics

Using the following metrics, the efficiency of the SGAN-OECR model is compared to that of cutting-edge models: "Accuracy, Precision, Recall and F1-Score" in equations 17 – 20, respectively. The quantitative formulas used to calculate the metrics are as follows:

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \quad (17)$$

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (20)$$

From the above equations (17-20), True Positive (TP) appropriately identifies the proportion of positive samples. In contrast, True Negative (TN) correctly categorises the number of negative samples. False Positive (FP) refers to the proportion of negative class samples that are identified as positive class samples, whereas False Negative (FN) refers to

the proportion of positive class samples that are classed as negative class samples.

#### 4.3. Performance Evaluation

This section validates the suggested SGAN-OECR model using the "Accuracy, Precision, Recall, and F1-score" from the available benchmark databases. The SGAN-OECR model was trained on two distinct databases for classifying RF pictures into different degrees of DR severity. Both databases use multiclass classification since the database has five classes. In addition, the conclusions of the APTOS and IDRiD tests are described in the parts that follow, including the performance comparison table and confusion matrices. APTOS and IDRiD class-wise performance evaluations are offered. Table 3 displays the comparative performance analysis of the presented model and existing methods for DR categorisation using the Kaggle-APTOS 2019 Blindness Detection dataset.

Table 3 combines and evaluates the modern framework of the Kaggle-APTOS database. According to the research, CNN is the most extensively used deep learning model for clinical image assessment. Based on the "Accuracy, Precision, Recall, and F1-Score", Table 3 evaluates the efficacy of the SGAN-OECR model on the Kaggle-APTOS database using four categorisation models from current research. According to Table 3, the suggested model surpasses cutting-edge Kaggle-APTOS models.

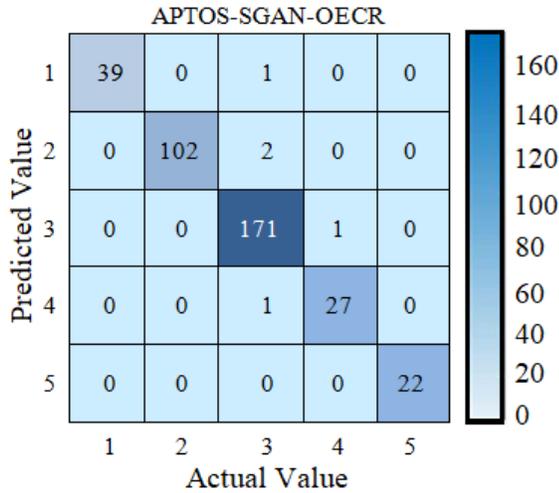


Fig. 8 Confusion matrix for SGAN-OECR model on the Kaggle-APTOS 2019 blindness detection database

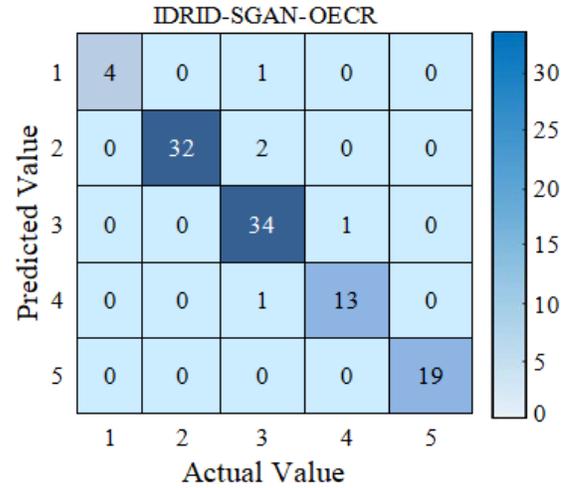


Fig. 9 Confusion matrix for SGAN-OECR model on the IDRiD

Existing models like ResNetGB [13], DT-EL [18], VGG-NIN [21], U-Net [22], Funswim [26], DCNN [27], MSA-ResNetGB [15], and SGAN-ECR achieved an accuracy of 88.40%, 86.81%, 87.65%, 87.37%, 88.89%, 89.47%, 94.40%, 97.81%. Thus, it is noticed that these models' performance is better but not superior to that of the suggested model. This validates the reliability of the suggested model.

Figure 8 depicts the confusion matrix for the suggested SGAN-OECR algorithm on the APTOS database. It provides the prevalence of images by category and the fraction of accurately and incorrectly classified images. In class 0, 39 images were accurately diagnosed as having no DR, 102 images were accurately categorised in class 1 as having mild DR, 171 images were accurately categorised in class 2 as having considerable DR, 27 images were accurately defined in class 3 as having extreme DR, and 22 images were accurately classed in class 4 as having PDR.

While using the IDRiD dataset, the suggested model is trained and evaluated to generate DR categorisation results. The IDRiD consists of 516 photos, of which 20% of annotated images are evaluated for testing. The multiclass model is validated by comparing it to contemporary creative works. The efficiency study of the proposed and current algorithms for DR categorisation on the IDRiD is presented in Table 4.

In Table 4, relevant categorisation methods such as ResNetGB [13], CBAM-UNet [25], MSA-ResNetGB [15], and SGAN-ECR were used to assess the results of the SGAN-OECR model on the IDRiD using the evaluation metrics like "Accuracy, Precision, Recall, and F1-Score".

The existing models such as the CNN model [14], Mixed Model [20], ResNetGB [13], DTL model [23], CBAM-UNet [25], HDL Model [28], MSA-ResNetGB [15], and SGAN-ECR [16] obtained 86.02%, 87.50%, 88.89%, 89.01%, 89.47%, 90.11%, 94.17%, 96.12%, 99.03% accuracy, respectively, which is 15.12%, 13.17%, 11.41%, 11.25%, 10.69%, 9.89%, 5.17% and 3.03% less than the proposed model.

Figure 9 depicts the confusion matrix for the SGAN-OECR model on the IDRiD. Figure 9 depicts the distribution of photos by class and the fraction of correctly and incorrectly classified photographs. In class 0, 4 photos were appropriately recognised as having no DR; in class 1, 32 images were precisely categorised as having mild DR; in class 2, 34 images were accurately classified as having considerable DR; in class 3, 13 images were accurately defined as having extreme DR; and in class 4, 19 images were accurately categorised as having PDR.

The proper parameter selection for the classifier changes the model's architecture towards less error rate, as depicted in Figure 10. Figure 10 provides the average error between the restored and original images. The lower the Mean Square Error (MSE) value, the less degraded the restored image. This precaution prevents overfitting and, as a result, improves training effectiveness by drastically reducing the number of parameters generated during the training process. An optimal variable selection accelerates model training and convergence.

On the basis of attempting specific hyper-parameters, the produced model is used to estimate the value of all hyper-parameters space. Then, a specific number of high-quality hyper-parameters are picked to train the model repeatedly.

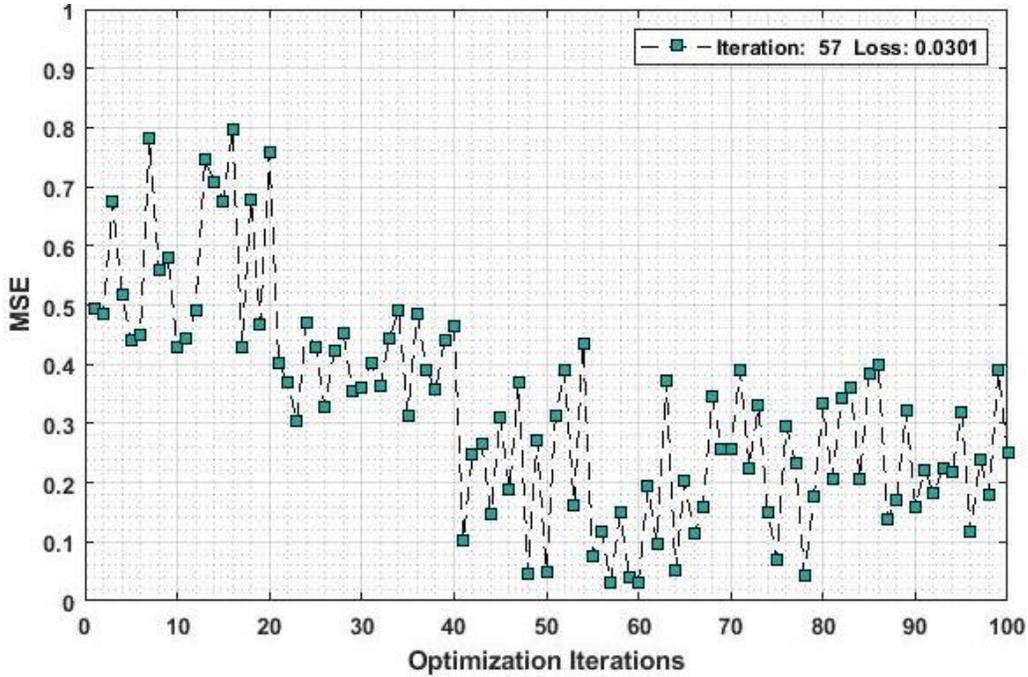


Fig. 10 MSE vs optimisation iterations

Table 5. Configuration of hyper-parameters for MSA-ResNet and GAN structure

Model	Hyper-parameter	Range	Optimal Value
MSA-ResNet	Batch Size	[5, 10, 20]	20
	Learning Rate	[0.001, 0.01, 0.1, 0.2]	0.01
	Filter Size	[1-10]	5
	Number of Feature Map	[32-256]	128
	Number of Filters	50-500	400
	Drop out	0.1~0.2	0.2
	Channel Size	[18-30]	24
	Optimizer	[Adam, SGD, Adagrad, Adadelta]	Adam
	Activation Function	[“ReLU”, “tanh”, “Linear”, “Sigmoid”]	Sigmoid
	Pooling	[Adaptive, Global]	Adaptive
	Residual Blocks	[3, 8, 12, 3]	8
	Scaling	3-5	5
GAN	Generator (Gen_layersize)	[16-40]	32
	Discriminator [Disc_layersize]	[6-20]	18
	Batch Size	[32-128]	64
	Gen_learning rate	[1-50]	24
	Disc_learning rate	[1-50]	24

Table 6. Performance of SGAN-OECR model on the "Kaggle-APTOS and IDRiD database"

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
APTOS	99.03	99.39	96.00	97.47
IDRiD	98.63	98.83	98.28	98.55

The amount of training data diminishes steadily with each iteration, but the model's accuracy increases fast, and eventually, the suitable hyper-parameter is acquired. Table 5 depicts the Hyper-parameter and its range of MSA-ResNet and GAN. The suggested SGAN-OECR model is discussed, and its experiential values on the "Kaggle-APTOS and IDRiD

datasets" are evaluated. Using the classical models discussed above, the performance of the models on the two databases was evaluated separately. Both reference databases are proposed models that provide better results than the models in the literature. Table 6 shows results from tests on the "Kaggle-APTOS and IDRiD databases" using the SGAN-OECR

model. On the Kaggle-APTOS database, the SGAN-OECR model obtained 99.03% accuracy, 99.39% precision, 96.00% recall, and 97.47% F1-Score, outperforming the IDRiD database, which obtained 98.63% accuracy, 98.83% precision, 98.28% recall, and 98.55% F1-Score.

## 5. Conclusion

This article presented an automated method called SGAN-OECR model to enhance the efficiency of categorising the DR severity grades. At first, the binary encoding of MSA-ResNet architectures was computed. Each layer in the CNN module has its own set of hyper-parameters encoded, with all

layers having the same constant length dependent on layer count and input image size. Then, the metaheuristic algorithm (EMBA) was used to select the best hyper-parameters and improves the MSA-ResNet structure. Finally, the fitness function was calculated to improve the model's detection of DR lesion abnormalities in FIR. This model efficiently resolves the problem of designing an automated hyper-parameter of MSA-ResNet architectures for optimising DR lesion grading. At last, the experimental findings of SGAN-OECR on Kaggle-APTOS and IDRiD proved that it has 99.03% and 98.63% accuracy, correspondingly, compared to the existing DR grade classification models.

## References

- [1] Dóra Júlia Eszes et al., "Diabetic Retinopathy Screening in Patients with Diabetes Using a Handheld Fundus Camera: The Experience from the South-Eastern Region in Hungary," *Journal of Diabetes Research*, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Solomon and Goldberg, "ETDRS Grading of Diabetic Retinopathy: Still The Gold Standard?," *Ophthalmic Research*, vol. 62, no. 4, pp. 190-195, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Wei Zhou, Chengdong Wu, and Xiaosheng Yu, "Computer-Aided Diagnosis for Diabetic Retinopathy Based on Fundus Image," *2018 37th Chinese Control Conference (CCC), IEEE*, pp. 9214-9219, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Meteb Altaf et al., "Monitoring Diabetic Foot by Designing a New Smart Sole," *SSRG International Journal of Pharmacy and Biomedical Engineering*, vol. 7, no. 3, pp. 15-20, 2020. [[CrossRef](#)] [[Publisher Link](#)]
- [5] Nabila Eladawi et al., "Computer-Aided Diagnosis System Based on a Comprehensive Local Features Analysis for Early Diabetic Retinopathy Detection using OCTA," *Diabetes and Fundus OCT*, pp. 1-23, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] K.S. Sreejini, and V.K. Govindan, "Retrieval of Pathological Retina Images Using Bag of Visual Words and Plsa Model," *Engineering Science and Technology, An International Journal*, vol. 22, no. 3, pp. 777-785, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Gwenolé Quéllec et al., "Automatic Detection of Referral Patients due to Retinal Pathologies through Data Mining," *Medical Image Analysis*, 29, pp 47-64, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Ping-Nan Chen et al., "General Deep Learning Model for Detecting Diabetic Retinopathy," *BMC Bioinformatics*, vol. 22, no. 5, pp. 1-15, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Zhentao Gao et al., "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks," *IEEE Access*, vol. 7, pp. 3360-3370, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Ms.S.Deepa, and Mr.S.Vijayprasath, "Certain Investigation of the Retinal Hemorrhage Detection in Fundus Images," *SSRG International Journal of Electronics and Communication Engineering*, vol. 2, no. 2, pp. 24-34, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Mustapha Aatila et al., "Diabetic Retinopathy Classification Using Resnet50 and VGG-16 Pretrained Networks," *International Journal of Computer Engineering and Data Science*, vol. 1, no. 1, pp. 1-7, 2021. [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Yunpeng Chen et al., "Dual Path Networks," *Advances in Neural Information Processing Systems*, vol. 30, 2017. [[Google Scholar](#)]
- [13] Fahman Saeed, Muhammad Hussain, and Hatim A. Aboalsamh, "Automatic Diabetic Retinopathy Diagnosis using Adaptive Fine-Tuned Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 41344-41359, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] S.Bibiana Vincy, and Dr. C. Chitra, "Detection of Exudates in Color Fundus Image," *SSRG International Journal of Electronics and Communication Engineering*, vol. 2, no. 11, pp. 1-10, 2015. [[CrossRef](#)] [[Publisher Link](#)]
- [15] Valarmathi Srinivasan, and Vijayabhanu Rajagopal, "Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification," *International Journal of Intelligent Engineering and System*, vol. 15, no. 4, pp. 489-498, 2022. [[CrossRef](#)] [[Google Scholar](#)]
- [16] Jordi de la Torre, Aida Valls, and Domenec Puig, "A Deep Learning Interpretable Classifier for Diabetic Retinopathy Disease Grading," *Neurocomputing*, vol. 396, pp. 465-476, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] D. Jude Hemanth, Omer Deperlioglu, and Utku Kose, "An Enhanced Diabetic Retinopathy Detection and Classification Approach Using Deep Convolutional Neural Network," *Neural Computing and Applications*, vol. 32, no. 3, pp. 707-721, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Niloy Sikder et al., "Severity Classification of Diabetic Retinopathy Using an Ensemble Learning Algorithm through Analysing Retinal Images," *Symmetry*, vol. 13, no. 4, pp. 1-26, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Mohammed Al-Mijalli, "Hemodynamic in Human Diabetic Abdominal Aneurysmal Aorta Using Computational Fluid Dynamics," *SSRG International Journal of Pharmacy and Biomedical Engineering*, vol. 8, no. 2, pp. 1-5, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [20] Anas Bilal et al., "Diabetic Retinopathy Detection and Classification Using Mixed Models for a Disease Grading Database," *IEEE Access*, vol. 9, pp. 23544-23553, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Zubair Khan et al., "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408-61416, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Eman Abdelmaksoud et al., "Automatic Diabetic Retinopathy Grading System Based on Detecting Multiple Retinal Lesions," *IEEE Access*, vol. 9, pp. 15939-15960, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Hassan Tariq et al., "Performance Analysis of Deep-Neural-Network-Based Automatic Diagnosis of Diabetic Retinopathy," *Sensors*, vol. 22, no. 1, pp. 1-15, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] B. Venkaiahppalaswamy, PVGD Prasad Reddy, and Suresh Batha, "An Effective Diagnosis of Diabetic Retinopathy Based on 3d Hybrid Squeezenet Architecture," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, pp. 147-159, 2022. [[CrossRef](#)] [[Publisher Link](#)]
- [25] Alnur Alimanov, and Md Baharul Islam, "Retinal Image Restoration and Vessel Segmentation using Modified Cycle-CBAM and CBAM-Unet," *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, *IEEE*, pp. 1-6, 2022. [[Google Scholar](#)]
- [26] Zhaomin Yao et al., "Funswin: A Deep Learning Method to Analysis Diabetic Retinopathy Grade and Macular Edema Risk Based on Fundus Images," *Frontiers in physiology*, 1462, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Najib Ullah et al., "Diabetic Retinopathy Detection Using Genetic Algorithm-Based CNN Features and Error Correction Output Code SVM Framework Classification Model," *Wireless Communications and Mobile Computing*, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Muhammad Mohsin Butt et al., "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features," *Diagnostics*, vol. 12, no. 7, p. 1607, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] APTOS 2019 Blindness Detection Dataset. [Online]. Available: <https://www.kaggle.com/c/aptos2019-blindness-detection/>
- [30] Prasanna Porwal et al., "Indian Diabetic Retinopathy Image Dataset (Idrid): A Database for Diabetic Retinopathy Screening Research," *Data*, vol. 3, no. 3, p. 25, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]