Journal of Autonomous Intelligence (2024) Volume 7 Issue 3
doi: 10.32629/jai.v7i3.1095

ORIGINAL RESEARCH ARTICLE

LW-CNN-based extraction with optimized encoder-decoder model for detection of diabetic retinopathy

B. Gunapriya[1], T. Rajesh[2], Arunadevi Thirumalraj[3,*], Manjunatha B[1]

[1] *New Horizon College of Engineering, Bengaluru 560103, India*

[2] *Electrical and Electronics Engineering Department, Malla Reddy Engineering College, Hyderabad 500100, India*

[3] *Department of Computer Science Engineering, K.Ramakrishnan College of Technology, Trichy 621112, India*

*** Corresponding author:** Arunadevi Thirumalraj, aruna.devi96@gmail.com

ABSTRACT

In the field of computer vision, automatic diabetic retinopathy (D.R.) screening is a well-established topic of study. It's tough since the retinal vessels are hardly distinguishable from the backdrop in the fundus picture, and the structure is complicated. To learn data representations at numerous levels of abstraction, deep learning (DL) allows for the development of computational representations with several processing layers. Small, inconspicuous lesions generated by the disorder are hard to detect since they are tucked away beneath the eye's structure. In this research, a lightweight convolutional neural network (LW-CNN) was used to extract structures from images of blood vessels, and different preprocessing methods were employed. The features are extracted, and then D.R. is classified using the suggested learning technique, which includes an encoder, dense branch. Effective categorization relies on the usage of multi-scale information collected from various nodes in the network. Grasshopper's optimisation algorithm (GHOA) is used to fine-tune the recommended classifier's hyper-parameters. The DIARETDB1 benchmark dataset is assessed using 80% training data and 20% testing data to get a diagnosis of the disease's severity. The proposed model improved D.R. image classification with accuracy of 0.992 for DIARETDB1 database and 0.981 for APTOS 2019 blindness detection dataset. The state-of-the-art models for D.R. dataset images only achieved less accuracy and precision as compared with the proposed model.

Keywords: diabetic retinopathy; lightweight convolutional neural network; grasshopper optimization algorithm; encoder structure; blood vessel feature extraction; fundus image

ARTICLE INFO

Received: 5 August 2023
Accepted: 4 December 2023
Available online: 29 December 2023

1. Introduction

Hyperglycemia, a chronic rise in blood glucose levels, is a hallmark of a group of metabolic illnesses related to diabetes mellitus (DM)[1,2]. Clinically, there are distinctions between type 1 and type 2 diabetes in terms of symptoms, diagnosis, comorbidities, and modes of treatment[3]. Diabetes has both immediate and long-term consequences, such as hyperosmolar hyperglycemic states, diabetic ketoacidosis, nerve damage, ocular impairments, chronic kidney disease, cardiovascular disease, stroke, and foot ulcers[4–6]. A major microvascular complication of diabetes mellitus, diabetic retinopathy (DR) is the primary cause of visual impairment in adults. According to projections, up to 34% of people in the world who are 40 years of age or older may have diabetes by 2035, underscoring the increasing influence of DR on global health[7]. Gender disparities in type 2 diabetes patients' prevalence and severity of DR are highlighted by research by Lin et al.[8]. DR is linked to long QT (process between the beginning of the QRS complex and the end of the T wave) intervals, potentially lethal arrhythmias, and mortality from non-cancer and vascular causes in addition to a declining quality of life[9]. Proliferative diabetic retinopathy (PDR) and non-proliferative diabetic retinopathy (NPDR) are the two types of diabetic retinopathy (DR) classified clinically. In patients with proliferative diabetic retinopathy, hypoxia-induced neovascularization in PDR is a primary indication for vitrectomy because it causes complications like retinal detachment and vitreous hemorrhage[9]. In contrast, early vascular abnormalities in NPDR lead to retinal pathologies. The four stages of diabetic retinopathy mild, moderate, severe, and PDR are distinguished by their respective clinical characteristics and severity[10,11]. Furthermore, convolutional neural networks (CNN), which are essential to deep learning and are frequently employed in pattern recognition, are significant[12,13]. These networks work well for a variety of applications because they are made up of interconnected "neurons" arranged into multiple layers. The sample image of DR is given in **Figure 1**.

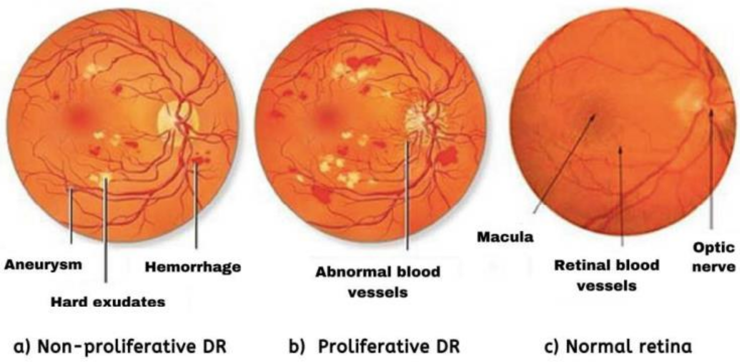

Figure 1. (a) and **(b)** Diabetic retinopathy cases; **(c)** usual retina.

It is becoming increasingly important to use fundus images for pathological screening in contemporary medical diagnosis. In recent years, automatic D.R. detection has received increasing interest from computer vision specialists. Since the results of the segmentation are so important to the success of the classification, several articles have been written on them[14]. Fundus pictures are collected in a variety of settings and at varying resolutions, making it difficult to standardise the data needed for an autonomous screening system. Most microaneurysms also present with bleeding[15]. Both NPDR and PDR are characterised by lesions of varied intensities, which raises the risk of consequences for the retinal arteries and optic disc. To address the improbable symmetry in connections across network layers, this work presents an asymmetric architecture[16] in which the weights are changed independently feedforward. In the arena of medical imaging, the two main responsibilities of a computerised system are classification and feature extraction. The image's features can be used to help diagnose D.R. and cure it early on. Therefore, it is an interesting path to investigate[17] to train a network to handle two goals (feature extraction and classification of D.R.) by exchanging features.

This work offers important new insights in a number of areas.

- First, it uses preprocessing methods to feed the preprocessed output into the extraction process, including data augmentation, color adjustments, and optic disc detection.
- Second, a unique feature extraction method called LW-CNN is suggested. Using a specially adapted design with a residual module, an encoder-decoder network is built for the classification of fundus images associated with Diabetic Retinopathy (DR).
- Finally, the GHOA model is used to hyper-tune the classifier models' parameters after the classification process.
- Together, these additions improve the methodology and efficacy of the suggested strategy for resolving fundus image analysis difficulties, especially when it comes to diabetic retinopathy.

The remainder of the paper is laid out as shadows: The current literature is discussed in Section 2, the materials and methods employed are outlined in Section 3, the reported results are discussed in Section 4, and the study is transported to a close in Section 5.

2. Related works

A novel encoder-decoder network architecture was introduced to improve the precision and effectiveness of microvascular lesion segmentation in a recent study by Yi et al.[18]. The training process was accelerated by using a lightweight encoder with more resolution, depth, and width. To fully utilize both spatial and channel-wise information, the decoder phase included Concurrent Spatial and Channel Squeeze and Channel Excitation (scSE) blocks in addition to an attention mechanism. The study addressed imbalanced data by using a compound loss function and transfer learning, which enhanced overall performance. Two large lesion segmentation datasets, FGADR and DDR, were used for evaluation, demonstrating the superiority of the proposed model over current methods.

A study on the semantic segmentation of exudates in retinal images using deep convolutional neural network (CNN) architectures was presented by Manan et al.[19]. The paper proposed a residual CNN with residual skip connections for robust and accurate exudate segmentation through efficient image augmentation techniques. To validate the effectiveness of the suggested method, a comparative performance analysis was carried out on benchmark databases, namely E-ophtha, DIARETDB1, and Hamilton Ophthalmology Institute's Macular Edema. The analysis focused on precision, accuracy, sensitivity, specificity and area under the curve.

A hybrid CNN+Transformer DR recognition and grading system was presented by Sadeghzadeh et al.[20]; even with training on small datasets, the system demonstrated competitive performance. The architecture included a deep CNN-based EfficientNet-B0 backbone as a feature extractor, a Transformer encoder-decoder with Multi-Head Self Attentions (MHSA), and a Residual Spatial Module (RSM). The approach demonstrated state-of-the-art performance in both grading and recognition on the EyePACS, APTOS, DDR, Messidor-l, and Messidor-2 benchmark datasets, demonstrating the synergies between CNN and Transformer architectures.

Chetoui et al.[21] used a federated learning strategy based on the Vision Transformer architecture to develop a robust model for diabetic retinopathy detection in collaboration with four institutions. In addition to improving accuracy by as much as 3%, this method addressed important concerns with data security, protection, and access rights. area under curve, SE, specificity, accuracy, and other performance metrics were taken into account in the study, which demonstrated how well the federated learning strategy handled data-related problems.

A unique Transformer-based model for DR segmentation that includes hyperbolic embeddings and a spatial prior module was presented by Wang et al.[22]. The model, which was constructed with a Vision Transformer encoder, performed better in terms of DR segmentation accuracy than other well-known models. Improved feature continuity and the model's capacity to capture underlying geometric structures were made possible by the spatial prior module and hyperbolic embeddings, which also improved automated DR diagnosis.

Wavelet and multi-Wavelet transforms combined with Swin Transformer are used to automatically identify stages of diabetic retinopathy progression, as reported by Dihin et al.[23]. Excellent training and test accuracies were made possible by the study's creative application of multi-Wavelet transforms for feature extraction, especially in binary classification tasks. But in multiclass classification, the model's accuracy decreased, suggesting that more research is needed in a variety of classification scenarios.

The Attention-DenseNet model, which incorporates an attention model to focus on specific areas for DR detection and grading, was proposed by Dinpajhouh et al.[24]. When tested on the APTOS 2019 dataset, the model performed better than other research and showed promise for DR classification in practical settings. For both the grading and detection tasks, it showed high accuracy, quadratic weighted kappa, and area under the receiver operating characteristic curve.

Using unlabeled data from Kaggle-EyePACS, Ullah et al.[25] presented a semi-supervised multitask learning technique for enhancing DR segmentation performance. After a thorough evaluation on publicly available datasets (FGADR and IDRiD), the model—which has a novel multi-decoder architecture surpassed state-of-the-art methods, demonstrating improved cross-data evaluation robustness and generalization.

Research gap

Although diabetic retinopathy (DR) detection using different deep learning architectures and methodologies have advanced recently, there is still a significant research gap in understanding how robust these models are across different datasets and stages of lesion progression[26]. The majority of studies conducted today concentrate on particular datasets or lesion types, which restricts their applicability to a range of imaging conditions and stages of disease response. Furthermore, even though some models show encouraging results on benchmark datasets, factors like image quality, a variety of lesion manifestations, and varying degrees of lesion severity may cause these models' performance to differ significantly in real-world clinical settings[27,28]. Furthermore, a thorough comparative analysis across several DR grading scales or datasets is lacking, which is essential for evaluating how adaptable and scalable these models are for clinical deployment. In the current landscape of DR detection methodologies, which requires a thorough examination that methodically assesses the generalization abilities, scalability, and performance consistency, the proposed model could be effective in producing results strong across a variety of datasets, lesion types, and severity levels while taking real-world clinical variability into account.

3. Proposed system

3.1. Preprocessing techniques

3.1.1. Data augmentation

It might be challenging to enhance pattern recognition when medical pictures used for training are inadequate. In order to improve robustness and generalizability, data development is required to artificially expand the extent of the dataset. The study generated an extra 1552 synthetic data with 3662 images from APTOS 2019 blindness detection dataset by applying geometric changes such as flips (vertical and horizontal) and rotations. The DIARETDB1 dataset contain only limited number of images so those images were not taken for this process. The model's predictive accuracy and resilience can both see improvements through this procedure. Finally, the paper used a 10-fold to separate the dataset into 80% for training purpose and for testing set is 20%, while also excluding ten early case D.R. instances.

3.1.2. Color adjustments for extreme feature extraction

Colour calibration is essential in image processing because it enhances visual quality and is tailored to a specific use case. Using blue (B) = 0.17 (**Figure 2**), the study converts to grayscale to decrease data size and decrease computing time. Additionally, this improves test run training efficiency.

Figure 2. **(a)** Sample grey scaled fundus images, usual; **(b)** and **(c)** irregular.

This paper findings go counter to the common assumption that blood vessels and clots should be highlighted by placing a premium on the red channels. Instead, this paper used an additional alteration of R = 0.28, G = 0.57, and B = 0.15 to emphasise the green colours over the channels since they bring out more detail. **Figure 3** shows the relative importance of each colour change.

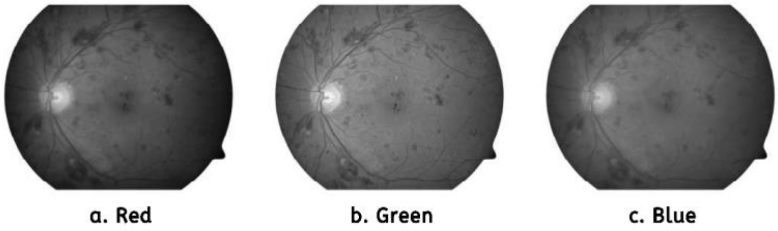

Figure 3. Ordering the green channel shaped finer and greater info **(b)** than the additional channels **(a)** and **(c)**.

3.2. Optic disc detection and elimination

Accurate diagnosis of retinal illnesses is achieved by careful inspection of the Optic Disc pixels in the fundus picture. The present study employs image processing techniques to find the retinal pictures by analysing the pixels with increased intensity, as per the proposed methodology. Grayscale pictures are used to determine where on the disc the optical drive is located. Using a median filter, the grayscale picture may be denoised once processing has begun[29]. Therefore, it was determined through analysis that the optical disc represents the pixels with increased pictures backgrounds. Therefore, after applying the filter, the associated intensity values are made uniform, but the pixels with high-intensity values stand out from the rest of the image. Filter range measurement is an integral part of settling on a suitable optic disk position. The group that exited with some pixels is responsible for figuring out where the optic disc pixels are, and this differs depending on whether the pixels came from the left or right eye. In the end, this synchronisation yields a segment in the form of a circular shape with a specified radius, commonly referred to as an optical disc. Last but not least, morphological processes are conducted to the binary picture, and the optic disc is extracted appropriately by erosion and dilation to refine and accurately identify the outside borders of the candidate disc pixels.

3.3. Blood vessels feature extraction using LW-CNN

While the red station becomes saturated with the value and the blue channel principals to noise and knowledge from unfavourable, the space exhibits the superior contrast in the retinal image. MobileNet, a lightweight deep neural network with a design, is used as the CNN model in this study since it has shown excellent classification performance. Among these are the fewer network parameters and cheaper computation expenses made possible by MobileNetV2's inverted residual and linear bottleneck topology. **Table 1** displays the configuration settings for the aforementioned network.

Table 1. Thorough parameters of each layer of the sum of repetitions of the Operator.

Input	Channel	Operator	Stride	N	Out
7 × 7 × 320	1280	conv2d	1	1	7 × 7 × 1280
7 × 7 × 1280	-	avgpool	-	1	1 × 1 × 1280
1 × 1 × 1280	3	conv2d	1	1	1 × 1 × 3
1 × 1 × 3	3	softmax	-	1	3
224 × 224 × 3	32	conv2d	2	1	112 × 112 × 32
112 × 112 × 32	16	bottleneck	1	1	112 × 112 × 16
112 × 112 × 16	24	bottleneck	2	2	56 × 56 × 24
56 × 56 × 24	32	bottleneck	2	3	28 × 28 × 32
28 × 28 × 32	64	bottleneck	2	4	14 × 14 × 64

Table 1. (*Continued*).

Input	Channel	Operator	Stride	N	Out
14 × 14 × 64	96	bottleneck	1	3	14 × 14 × 96
14 × 14 × 96	160	bottleneck	2	3	7 × 7 × 160
7 × 7 × 160	320	bottleneck	1	1	7 × 7 × 320

The MobileNetV2 network employs regular convolution (Conv), a deep separation convolution with an inverse residual structure (Bottleneck), and an average pooling layer (Avgpool). To broaden the network's use in classifying tea and raise the bar for target classification accuracy. The following enhancements have been added to MobileNetV2 for this investigation (as seen in **Figure 4** and **Table 2**). In order to spare even more CPU time and RAM during network training. Take the convolutional layer's channel count down from 1280 to 128, and get rid of the network layer that comes after the ninth. The Flatten layer flattened the original three-dimensional feature map. SoftMax was swapped out for SVM classifier to capabilities and make it more amenable to the challenge of tea hyperspectral picture classification.

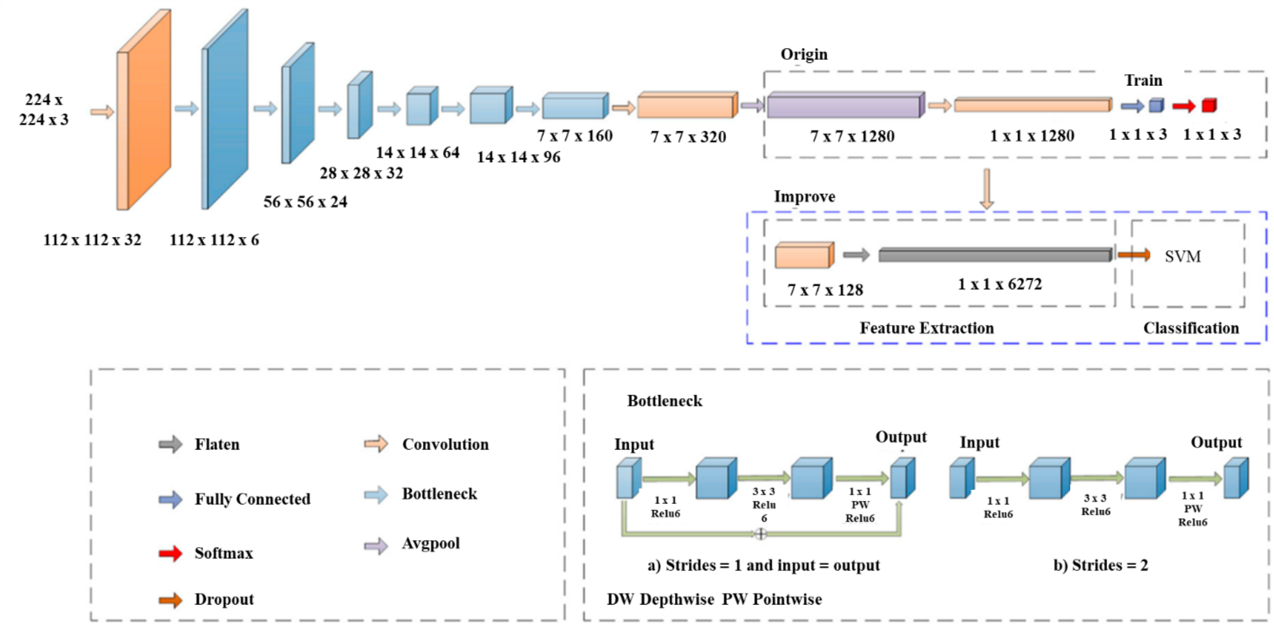

Figure 4. The enhanced lightweight CNN.

Table 2. Comprehensive parameters of CNN. (N: signifies the sum of repetitions of the Operator.)

Input	Channel	N	Stride	Operator	Out
14 × 14 × 96	160	3	2	bottleneck	7 × 7 × 160
7 × 7 × 160	320	1	1	bottleneck	7 × 7 × 320
224 × 224 × 3	32	1	2	conv2d	112 × 112 × 32
112 × 112 × 32	16	1	1	bottleneck	112 × 112 × 16
28 × 28 × 32	64	4	2	bottleneck	14 × 14 × 64
14 × 14 × 64	96	3	1	bottleneck	14 × 14 × 96
112 × 112 × 16	24	2	2	bottleneck	56 × 56 × 24
56 × 56 × 24	32	3	2	bottleneck	28 × 28 × 32
7 × 7 × 320	128	1	1	conv2d	7 × 7 × 128
7 × 7 × 128	6272	1	-	flatten	1 × 1 × 6272
1 × 1 × 6272	3	1	-	SVM	3

By using the LW-CNN model, the blood vessel features are extracted, which is shown in **Table 3**.

Table 3. Assessing features for precise credit of D.R. lesions.

Feature name	**Description**	**Feature category**
Eccentricity	Dissimilarity in extent is figured via quantifying elliptical values of an object by its chief axis.	Geometrical
Aspect ratio	Dissimilarity in dimension is evaluated over the ratio among the major axis and slight axis of the applicant area.	Geometrical
Mean	The regular charge of pixels (G-Channel) is limited by the appropriate D.R. part.	Statistical
Standard deviation	Measure the extent of spreading (variation) of the set of applicant pixels	Statistical
Variance	Difference in intensity value	Statistical
Area	The sum of pixels in the conceivable candidate area.	Geometrical
Perimeter	The difference in size is strongminded by the edge of the candidate district).	Geometrical
Solidity	The difference in object length is assessed by investigating the sum of holes.	Geometrical
Distance from optic disc	Pixel's site is the retinal copy.	Location dependent
Angle from optic disc	Way and site of applicant pixels in a copy.	Location dependent

Figure 5 presents the sample blood vessel features extraction using LW-CNN model.

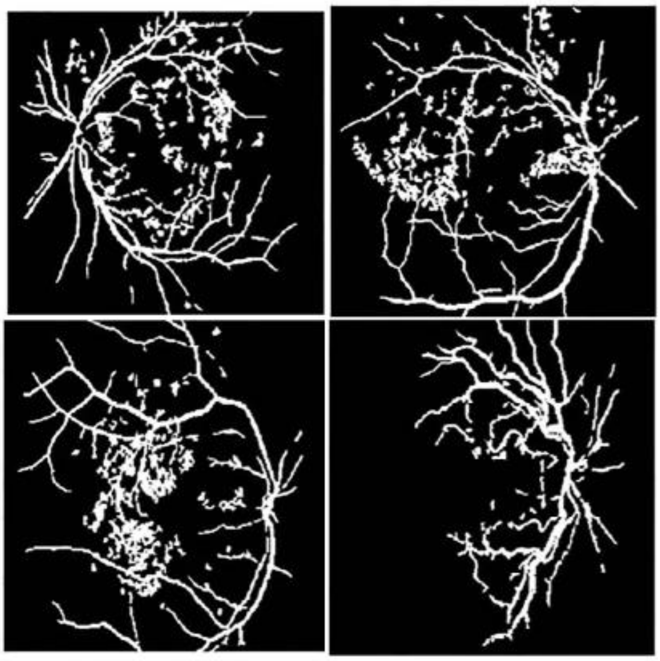

Figure 5. Sample Feature extraction.

3.4. Classification using optimized DL model

After the features are extracted, the classification process[30] should be done for accurate detection. The residual learning method was utilised in the classification job to facilitate the training of the networks and address the degradation issue. The layers employed in the proposed work were: 3 convolutions, 3 batch normalisation, 3 ReLU, and 1 concatenation. Each residual unit in the residual network comprises a batch normalisation layer and ReLU activation, and the remaining network is a stack of these remaining units. **Figure 6** compares the module used by normal UNet's neural layers with the suggested residual module.

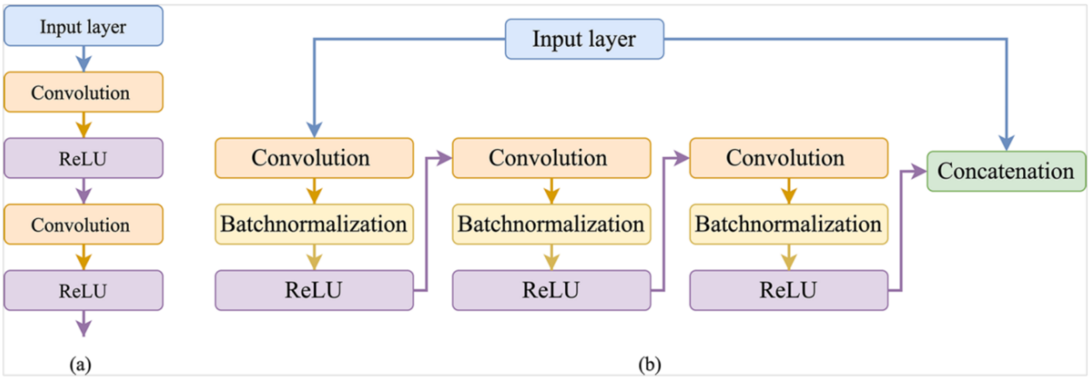

Figure 6. (a) The neural layers unit in traditional UNet and the outstanding unit suggested herein: U-Net's neural layers; **(b)** the residual module that was suggested.

Shortcut connections, also known as skip connections, are those that bypass whole neural network layers, as seen in **Figure 6a**. Equation (1) is a representation of the suggested residual unit seen in **Figure 6b**.

$$res_{out} = g_b\left(cv_{n\times n}\left(g_b(cv_{n\times n}(g_b(cv_{n\times n}((x_{in})))))\right)\right) + x_{in} \tag{1}$$

The batch-normalization layer is denoted by g_b in Equation (1), the convolutional layer by $cv_{n\times n}$, the input by x_{in}, and the residual block's output by res_{out}.

3.4.1. Residual-U-Net

In order to create the suggested network, the benefits of outstanding and U-Net topologies are combined. There are two benefits to combining these networks: (1) The residual unit facilitates the network's learning (2) The skip connection between the high and low levels in the residual unit will help for the propagation of information without degradation.

This study proposes a Residual-U-Net, a 9-layer design with a simple encoder, decoder, and bridge. The picture is compressed into a representation, which is then decoded by the decoder to yield pixel-by-pixel categorization. A bridge connects the encoder and decoder halves of a system. **Figure 6b** depicts the residual units used to construct these sections, which include three 3 × 3-filter convolutional layers, three 3 × 3-layers of batch normalisation, three 3 × 3-layers of ReLU activation, and a final identity mapping. In order to extract semantic information, the encoding block employs four down-sampling processes, one for each outstanding unit. Instead of employing a pool operation to maintain positional info, a stride of 2 is used to the first convolutional layer of each encoding residual unit to down sample by its half. The related decoder path has four residual units, with each unit's feature map being a unit and map from the path. The last encoding component is shadowed by a sigmoid activation layer and a 3x3 convolutional layer for picture projection.

The block is collected and fed into a single dense layer for categorization of the D.R. as benign or malignant. In contrast, the features employed to retrieved from the bridge and the first decoder block to do classification. The feature maps are extracted and combined in this stage. Because of their unique dimensions, these features require Global Average to be appended to the end of each block before they can be combined. To facilitate feature concatenation across several scales, GAP layers are used. Although both Max Pooling (GMP) layers do the same thing, each has its drawbacks. Training takes longer, and there are more factors to worry about when using F.C. layers. The GMP, on the other hand, utilises a max-vocal representation of all characteristics, which ignores a great deal of informative geographical data. Subsequently, these characteristics are subsequently transmitted to a succession of dense classifications. To forecast whether an input ultrasound picture is benign or malignant, which gets with the ReLU function, while the last layer, which has 2 units, is active with the function. The suggested classification tree comprises two thick layers, which is an improvement over Wang et al.[31]. To prevent overfitting, a dropout layer is introduced as regularisation to the network between these two dense layers, with a dropout rate of 0.5.

3.4.2. Loss function

Class imbalance is the most difficult challenge in medical imaging because it might cause the model to be too biased towards one class if it is not fixed. Since there are more malignant instances than benign cases in the datasets analysed here, the resulting model may be biased towards the former. Equation (2) depicts the usage of weighted focus loss for the classification job as a means of resolving this inequity.

$$Cls_{loss} = -w_m P_{class}^{\gamma}(1 - Y_{class})log(1 - P_{class}) - w_b(1 - P_{class})^{\gamma} Y_{class} log(P_{class}) \tag{2}$$

In the above equation, P_{class} is the foretold organization production, Y_{class} is the actual lesson, γ is wm and stand for the weights allocated to the benign classifications, respectively, and are equal to 2, the focusing parameter that can be predicted by GHOA. The relative values are displayed in Equation (3).

$$w_m = \frac{1}{N_m}, w_b = \frac{1}{N_b} \tag{3}$$

3.4.3. Hyperparameter tuning using GHOA

After the classification process, the tuning process can be done for getting an impressive result in diagnosing D.R. effectively. Hence for tuning the hyperparameters of classifier, this study introduces an optimisation model[32] to determine the value of the focusing parameter indicated in Equation (2). The suggested recognition model benefits from this as it becomes more precise and has a lower mistake rate. The challenges of real-time application structure optimisation are addressed by using GHOA[33]. It efficiently finds global solutions to local optimisation issues, returns relevant output, and probes uncharted regions of the search landscape. As a bonus, it makes optimal more precise. It is formed from the swarming behaviour of adults and nymphs. When the grasshopper is a nymph, it moves very slowly and in little bursts, but when it is an adult, it moves much farther and much quicker. In Equation (4), a simulation of the grasshopper's swarming behaviour is described.

$$Gr_i = So_i + Gf_i + Wa_i \tag{4}$$

The variables in Equation (4) are the *i*-th grasshopper's location, its social interactions, its advection by the wind, and the force of gravity on the *i*-th grasshopper. Gr_i, So_i, Wa_i and Gf_i, correspondingly. The swarming behaviour is adapted in Equation (5) by replacing the standards of So_i and Gf_i.

$$Gr_i = \sum_{j=1, j \neq i}^{N} so(|gr_j - gr_i|)\frac{gr_j - gr_i}{d_{ij}} - C_{\hat{e}_C} + D_{\hat{e}_w} \tag{5}$$

Here, $So_i = \sum_{j=1,j\neq1}^{N} so(|gr_j - gr_i|)\frac{gr_j - gr_i}{d_{ij}}$, where the distance amongst the *i*-th grasshopper and *j*-th grasshopper, where $d_{ij} = |gr_j - gr_i|$ and the sum of grasshoppers is called as N, and the term $so()$ describes forces. $Gf_i = -C_{\hat{e}_C}$, where $\hat{e}_C$ and C mean a unity of the constant, correspondingly. $Wa_i = D_{\hat{e}_w}$, in which $\hat{e}_w$ and D is a unity drift. The optimisation difficulties posed by Equation (6) are solved by modifying Equation (5).

$$Gr_i^x = dc\left(\sum_{j=1, j\neq1}^{N} dc\frac{up_x - lo_x}{2} s(|gr_j^r - gr_i^r|)\frac{gr_j - gr_i}{gr_{ij}}\right) + \hat{T}A_x \tag{6}$$

In Equation (6), a lessening coefficient is assumed as dc that is expressed in Equation (7), $\hat{T}A_x$ is the value of the target at *X*-th dimension, and the bound in the *X*-th dimension are signified as up_x and lo_x, correspondingly.

$$dc = dc\,max - it\frac{dc\,max - dc\,min}{IT} \tag{7}$$

Maximum and lowest dc values are given by dc max and dc min in Equation (7), whereas I.T. and it signify the maximum and current iterations, respectively. The values of dc are set at 1 and 0.00001, respectively. Algorithm 1 is a pseudo-code representation of the GHOA algorithm.

Algorithm 1 Grasshopper Optimization Algorithm (GHOA)

```
BEGIN
Initialization of grasshopper swarm Gri (i = 1, 2, 3, …, n);
Initialization of variables;
Compute the fitness of every search solution agent;
Consider the best search solution agent as TA;
While (it < IT) do
Update dc by Eq. (7);
For each (search solution agent) do
Normalization of distances between the grasshoppers;
Update the location of the current search agent solution using Eq. (6);
end for
Update if there is the best search solution agent;
it = it + 1;
end while
Return TA;
END.
```

4. Results and discussion

4.1. Dataset description

(a) Diabetic Retinopathy Data Quality Measurement Level 1 (DIARETDB1): In the context of D.R. recognition from digital photos, it is a reference to a publicly available benchmark database. The total sum of photos in the 'DIARETDB1 database' is 89, with NPDR collections represented by 84 images. A fundus captures these images of the retina[34]. Newly added images to the database often included some level of image noise. Images in this collection have a depth of 24 bits and a bit depth of 1152 × 1500 pixels and are stored in the portable network graphics (PNG) format[35]. There is a possibility that the database stores the truth that may be used to set up a variety of D.R. image anomalies. Retinal pictures from the gold-standard D.R. database is displayed in **Figure 7**.

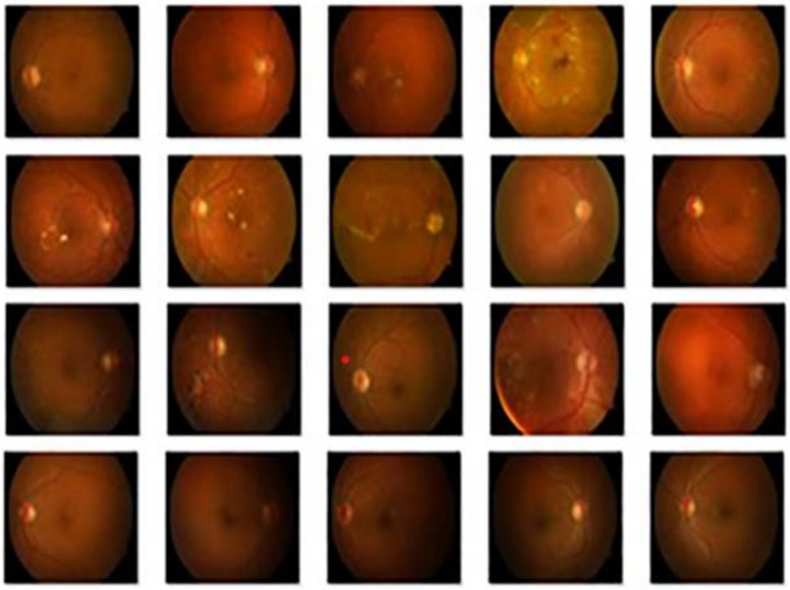

Figure 7. Example Images of the dataset.

(b) The APTOS dataset, which includes 3662 retinal images taken in various lighting conditions, was a major source of data for this study. These photos were taken from the Aravind Eye Hospital's APTOS 2019 Blindness Detection Database in India[36], which gave the study significant practical significance. Based on diabetic retinopathy (DR), the dataset categorized retinal images into five severity classes: Class 0 (non-DR), Class 1 (mild DR), Class 2 (moderate DR), Class 3 (severe DR), and Class 4 (proliferative DR), as shown in **Table 4** with samples. The dataset was carefully balanced in order to guarantee a balanced representation, which produced an equal image representation for every severity level. Effective model training and evaluation were made possible by this rigorous balancing process, which highlights the significance of balance in preserving the model's generalizability across various stages of DR.

Table 4. Level of severity.

Class	**0**	**1**	**2**	**3**	**4**
Classification	Non-DR	Mild DR	Moderate DR	Severe DR	Proliferative DR
Number of samples	1805	370	999	193	295

4.2. Experimental arrangement

The requested work has been implemented on a processing machine, with MATLAB R2020a installed for use. The computer has a clock speed of 2.5 GHz, runs on a 64-bit operating system, and has 4 G.B. of RAM.

For medical image processing, MATLAB R2020a has several advantages over other languages.

i This setting facilitates data exploration, enables persistent file tracking, and provides context for debugging.

ii As a result, it is extensively employed as a means of decoding both generic and specialized picture formats.

iii Thirdly, MATLAB R2020a can recognize third-party libraries like OpenCV.

iv Fourth, a MATLAB coder can automatically produce the code for you. It's a big library full of arithmetic and image processing operations, useful in many different contexts (embedded systems, the web, etc.).

4.3. Evaluation metrics

This study measured the $accuracy, precision, sensitivity, and\ specificity$ of the deep learning[37] architecture for classification purposes. Equations (8)–(11) outline typical machine learning standards using the terms including True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) as shown below.

$$Accuracy(AC) = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision(PR) = \frac{TP}{TP + FP} \tag{9}$$

$$Sensitivity(SE) = \frac{TP}{TP + FN} \tag{10}$$

$$Specificity(SP) = \frac{TN}{TN + FP} \tag{11}$$

The **Figures 8** and **9** depict the graphical format of the training and testing accuracy as well as loss for the APTOS dataset.

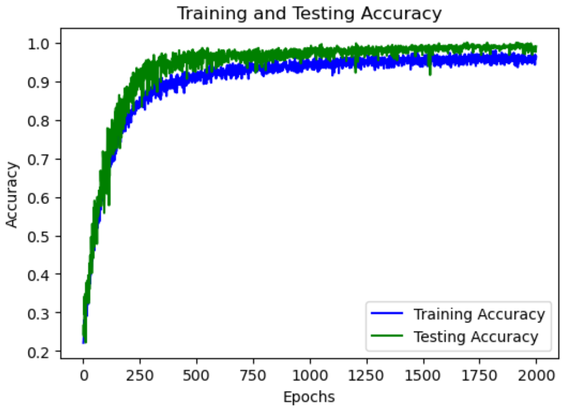

Figure 8. Illustration of training and testing accuracy.

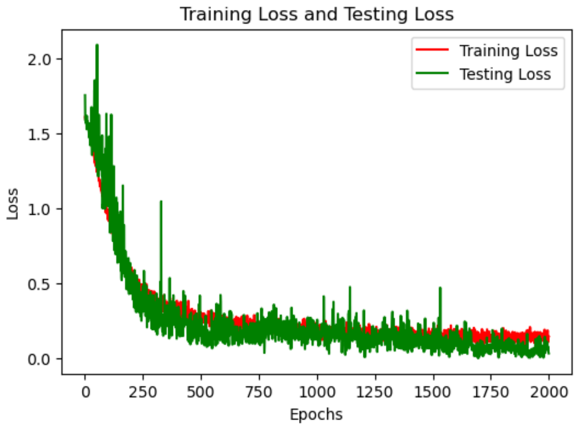

Figure 9. Illustration of training and testing loss.

4.4. Analysis of proposed classifier on two datasets

In this section, analysis of classification is done with two different classes of DIARETDB1 database. The penalties are given in **Tables 5** and **6**.

Table 5. Classification analysis on DIARETDB1 database.

Class	Recall %	F1-Score %	Specificity %	Precision %	Accuracy %
1 (Non-DR)	93.65	91.16	90.06	88.80	91.90
2 (Proliferative)	88.72	85.64	95.53	82.78	95.40
Macro Avg	87.04	88.23	95.67	89.86	94.80

In DIARETDB1 database, there are 2 different classes. In the classification analysis on DIARETDB1 dataset, the class 1 achieved an accuracy of 91.90% and class 2 gained an accuracy of 95.40%. The results in the above table proved that the classifier performed efficiently as both classes achieved best accuracy result. Hence, the macro avg calculation gained high accuracy level.

Table 6. Classification analysis on APTOS 2019 blindness detection dataset.

Class	Recall %	f1-Score %	Specificity %	Precision %	Accuracy %
0	99.35	98.50	97.95	97.66	98.60
1	87.00	89.69	99.22	92.55	98.00
2	91.81	92.30	98.08	92.84	97.25
3	92.08	92.54	96.99	93.00	95.50
4	88.81	88.48	98.15	88.15	96.90
Macro Avg	91.81	92.30	98.07	92.84	97.25

In APTOS dataset, there are 5 classes present. In the classification analysis on APTOS dataset, the class 0 achieved an accuracy rate of 98.60%, which gained highest ACC when compared with other classes from 1 to 4. The other classes also achieved better results in the analysis with proposed classifier. The robustness and efficient of the proposed model has been proved from the table metric values. Also, the macro avg calculation gained optimal ACC rate in diagnosing D.R.

4.5. Comparative analysis with existing works

The existing techniques such as CNN[19,21,28], U-Net[20], SqueezeNet[22], CNN+SVM[23], DELM[24], LSTM[25] and ELM[26] are considered for the experimental evaluation. But the above techniques use various datasets for D.R. classification. Hence it is verified with the datasets used in this paper, and then, consequences are around in **Table 7**.

Table 7. Comparative analysis of projected model with existing works in terms of various metrics on DIARETDB1.

Models	Accuracy	Precision	Sensitivity	Specificity
ELM	0.861	0.790	0.840	0.820
DELM	0.869	0.803	0.881	0.855
CNN	0.920	0.820	0.88	0.831
CNN+SVM	0.961	0.814	0.978	0.921
LSTM	0.966	0.953	0.953	0.973
SqueezeNet	0.985	0.962	0.996	0.982
U-Net	0.987	0.972	0.996	0.982
Proposed Model	**0.992**	**0.989**	**0.994**	**0.989**

When compared to the other models in the table mentioned above, the proposed model performed better on every metric averaged with two classes. The results in **Table 7** are not consistent with performance in **Table 5** as the DIARETDB1 contains only two classes including Non D.R. and Proliferative D.R. It outperformed well-known models like ELM, DELM, CNN, CNN+SVM, LSTM, SqueezeNet, and U-Net with accuracy reaching an astounding 99.2%, PR at 98.9%, sensitivity hitting 99.4%, and specificity achieving 98.9%. A number of factors, such as the application of cutting-edge methods, creative architectural design, or more efficient use of the underlying data, could be responsible for this outstanding performance. A balanced ability to accurately identify positive instances while minimizing false positives and false negatives is suggested by the high precision and sensitivity. Compared to other models, the suggested model’s accuracy and robustness make it an attractive and practical option for the particular task, demonstrating its potential for practical use in healthcare.

From **Table 8**, the proposed classifier model has the highest values for every metric, making it stand out as the best performer among the wide range of models examined. With a 98.1% accuracy rate, 97.8% precision rate, 98.6% sensitivity rate, and 97.9% rate, the suggested model strikes an amazing balance in accurate classification. A number of elements, such as the application of cutting-edge techniques, creative architectural design, and efficient use of the underlying dataset, are responsible for this better performance. In the context of the assessed metrics averaged with all classes, these results performance are consistent with the results in **Table 6**. Hence, the proposed model as a strong and dependable solution for the particular task, highlighting its potential for real-world applications and highlighting its superiority over well-known models like ELM, DELM, CNN, CNN+SVM, LSTM, SqueezeNet, and U-Net.

Table 8. Comparative analysis of projected model with existing works in terms of various metrics on APTOS 2019 blindness detection dataset.

Models	Accuracy	Precision	Sensitivity	Specificity
ELM	0.842	0.779	0.832	0.809
DELM	0.850	0.790	0.863	0.835
CNN	0.909	0.802	0.871	0.822
CNN+SVM	0.942	0.799	0.957	0.911
LSTM	0.945	0.933	0.944	0.957
SqueezeNet	0.964	0.954	0.971	0.965
U-Net	0.968	0.962	0.971	0.978
Proposed Model	**0.981**	**0.978**	**0.986**	**0.979**

5. Conclusion and future work

Globally, D.R. is caused by long-term diabetes, which leads to permanent vision loss. Manual identification of D.R. is not only labor-intensive but also needs PR in classifying lesions shown in retinal imaging. This necessitates the deployment of a computer-automated D.R. diagnostic in order to ascertain the degree of illness severity. From the DIARETDB1 dataset, this method offers segmentation of the optic disk, vessels, and classification. To begin, the methodical foundation for subsequent phases is provided by the preprocessing stage's translation of RGB colour space via the scaling procedure. Second, the filtered image's high-intensity pixels have been identified, and a candidate disc has been recovered after the edges have been refined, utilising dilation and erosion techniques for optic disc segmentation. The LW-CNN model is employed for blood vessel extraction, while the optimised DL model is used for classification. In this study, an organization branch is added to the U-Net architecture and residual units are used to construct the classifier model. For effective categorization of the D.R. phases, multi-scale characteristics are retrieved from various layers of the projected U-Net. The recommended model outdoes state-of-the-art classification techniques in terms of demonstrated by the experiments. This model improved D.R. image classification with accuracy of 0.992 for DIARETDB1 database and 0.981 for APTOS 2019 blindness detection dataset. This demonstrates the model's superiority over preexisting ones since it has reported fewer false negatives. A DR-based method will be suggested in the future to modify the CNN architecture parameters. The next proposed technique will determine the number of convolution and pooling layers, the number of kernels, and the size of the kernels of the convolution layer. Consequently, it is possible to decrease the number of untrainable hyperparameters. There are certain difficulties in converting DR to a CNN. For a clear classification, the maximum and minimum kernel sizes will be specified based on the dimensions of the input image.

Author contributions

Study conception and design, BG, TR; data collection, AT; analysis and interpretation of results, BG, TR, AT, MB; draft manuscript preparation, AT, MB. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Zhang J, Deng Y, Wan Y, et al. Diabetes duration and types of diabetes treatment in data-driven clusters of patients with diabetes. Frontiers in Endocrinology. 2022, 13. doi: 10.3389/fendo.2022.994836
2. Cole JB, Florez JC. Genetics of diabetes mellitus and diabetes complications. Nature Reviews Nephrology. 2020, 16(7): 377-390. doi: 10.1038/s41581-020-0278-5

3. Udler MS. Type 2 Diabetes: Multiple Genes, Multiple Diseases. Current Diabetes Reports. 2019, 19(8). doi: 10.1007/s11892-019-1169-7
4. Xiong X, Yang Y, Wei L, et al. Identification of two novel subgroups in patients with diabetes mellitus and their association with clinical outcomes: A two-step cluster analysis. Journal of Diabetes Investigation. 2021, 12(8): 1346-1358. doi: 10.1111/jdi.13494
5. Tanabe H, Saito H, Kudo A, et al. Factors Associated with Risk of Diabetic Complications in Novel Cluster-Based Diabetes Subgroups: A Japanese Retrospective Cohort Study. Journal of Clinical Medicine. 2020, 9(7): 2083. doi: 10.3390/jcm9072083
6. Zaharia OP, Strassburger K, Strom A, et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: A -year follow-up study. The Lancet Diabetes & Endocrinology. 2019, 7(9): 684-694. doi: 10.1016/s2213-8587(19)30187-1
7. Tong N, Wang L, Gong H, et al. Clinical Manifestations of Supra-Large Range Nonperfusion Area in Diabetic Retinopathy. International Journal of Clinical Practice. 2022, 2022: 1-7. doi: 10.1155/2022/8775641
8. Lin K, Hsih W, Lin Y, et al. Update in the epidemiology, risk factors, screening, and treatment of diabetic retinopathy. Journal of Diabetes Investigation. 2021, 12(8): 1322-1325. doi: 10.1111/jdi.13480
9. Takao T, Suka M, Yanagisawa H, et al. Combined effect of diabetic retinopathy and diabetic kidney disease on all-cause, cancer, vascular and non-cancer non-vascular mortality in patients with type 2 diabetes: A real-world longitudinal study. Journal of Diabetes Investigation. 2020, 11(5): 1170-1180. doi: 10.1111/jdi.13265
10. Gomułka K, Ruta M. The Role of Inflammation and Therapeutic Concepts in Diabetic Retinopathy—A Short Review. International Journal of Molecular Sciences. 2023, 24(2): 1024. doi: 10.3390/ijms24021024
11. Wang CY, Mukundan A, Liu YS, et al. Optical Identification of Diabetic Retinopathy Using Hyperspectral Imaging. Journal of Personalized Medicine. 2023, 13(6): 939. doi: 10.3390/jpm13060939
12. Serey J, Alfaro M, Fuertes G, et al. Pattern Recognition and Deep Learning Technologies, Enablers of Industry 4.0, and Their Role in Engineering Research. Symmetry. 2023, 15(2): 535. doi: 10.3390/sym15020535
13. Deshpande NM, Gite SS, Aluvalu R. Microscopic Analysis of Blood Cells for Disease Detection: A Review. Tracking and Preventing Diseases with Artificial Intelligence. 2021, 125-151. doi: 10.1007/978-3-030-76732-7_6
14. Gangwar AK, Ravi V. Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning. Advances in Intelligent Systems and Computing. 2020, 679-689. doi: 10.1007/978-981-15-5788-0_64
15. Yi SL, Yang XL, Wang TW, et al. Diabetic Retinopathy Diagnosis Based on RA-EfficientNet. Applied Sciences. 2021, 11(22): 11035. doi: 10.3390/app112211035
16. Liu H, Yue K, Cheng S, et al. Hybrid Model Structure for Diabetic Retinopathy Classification. Journal of Healthcare Engineering. 2020, 2020: 1-9. doi: 10.1155/2020/8840174
17. Das S, Kharbanda K, M S, et al. Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy. Biomedical Signal Processing and Control. 2021, 68: 102600. doi: 10.1016/j.bspc.2021.102600
18. Yi D, Baltov P, Hua Y, et al. Compound Scaling Encoder-Decoder (CoSED) Network for Diabetic Retinopathy Related Bio-marker Detection. IEEE Journal of Biomedical and Health Informatics. 2023, 1-12. doi: 10.1109/jbhi.2023.3313785
19. Manan MA, Jinchao F, Khan TM, et al. Semantic segmentation of retinal exudates using a residual encoder–decoder architecture in diabetic retinopathy. Microscopy Research and Technique. 2023, 86(11): 1443-1460. doi: 10.1002/jemt.24345
20. Sadeghzadeh A, Junayed MS, Aydin T, et al. Hybrid CNN+Transformer for Diabetic Retinopathy Recognition and Grading. 2023 Innovations in Intelligent Systems and Applications Conference (ASYU). 2023. doi: 10.1109/asyu58738.2023.10296789
21. Chetoui M, Akhloufi MA. Federated Learning for Diabetic Retinopathy Detection Using Vision Transformers. BioMedInformatics. 2023, 3(4): 948-961. doi: 10.3390/biomedinformatics3040058
22. Wang Z, Lu H, Yan H, et al. Vison Transformer Adapter-based Hyperbolic Embeddings for Multi-lesion Segmentation in Diabetic Retinopathy. 2023. doi: 10.21203/rs.3.rs-2728770/v1
23. Dihin RA, AlShemmary E, Al-Jawher W. Diabetic Retinopathy Classification Using Swin Transformer with Multi Wavelet. Journal of Kufa for Mathematics and Computer. 2023, 10(2): 167-172. doi: 10.31642/jokmc/2018/100225
24. Dinpajhouh M, Seyyedsalehi SA. Automated detecting and severity grading of diabetic retinopathy using transfer learning and attention mechanism. Neural Computing and Applications. 2023, 35(33): 23959-23971. doi: 10.1007/s00521-023-09001-1
25. Ullah Z, Usman M, Latif S, et al. SSMD-UNet: semi-supervised multi-task decoders network for diabetic retinopathy segmentation. Scientific Reports. 2023, 13(1). doi: 10.1038/s41598-023-36311-0
26. Nahiduzzaman Md, Robiul Islam Md, Omaer Faruq Goni Md, et al. Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier. Expert Systems with Applications. 2023, 217: 119557. doi: 10.1016/j.eswa.2023.119557
27. Alwakid G, Gouda W, Humayun M. Deep Learning-Based Prediction of Diabetic Retinopathy Using CLAHE and ESRGAN for Enhancement. Healthcare. 2023, 11(6): 863. doi: 10.3390/healthcare11060863

28. Ishtiaq U, Abdullah ERMF, Ishtiaque Z. A Hybrid Technique for Diabetic Retinopathy Detection Based on Ensemble-Optimized CNN and Texture Features. Diagnostics. 2023, 13(10): 1816. doi: 10.3390/diagnostics13101816
29. Sarathi MP, Dutta MK, Singh A, et al. Blood vessel inpainting based technique for efficient localization and segmentation of optic disc in digital fundus images. Biomedical Signal Processing and Control. 2016, 25: 108-117. doi: 10.1016/j.bspc.2015.10.012
30. Thirumalraj A, Asha V, Kavin BP. AI and IoT-Based Technologies for Precision Medicine. 2023
31. Wang P, Patel VM, Hacihaliloglu I. Simultaneous Segmentation and Classification of Bone Surfaces from Ultrasound Using a Multi-feature Guided CNN. Lecture Notes in Computer Science. 2018, 134-142. doi: 10.1007/978-3-030-00937-3_16
32. Das R M, Thirumalraj A, Rajesh T. An Improved ARO Model for Task Offloading in Vehicular Cloud Computing in VANET. 2023.
33. Saremi S, Mirjalili S, Lewis A. Grasshopper Optimisation Algorithm: Theory and application. Advances in Engineering Software. 2017, 105: 30-47. doi: 10.1016/j.advengsoft.2017.01.004
34. Kauppi T, Kalesnykiene V, Kamarainen JK, et al. the DIARETDB1 diabetic retinopathy database and evaluation protocol. Procedings of the British Machine Vision Conference 2007. 2007. doi: 10.5244/c.21.15
35. DIARETDB1. Diaretdb1 Diabetic Retinopathy Database and Evaluation Protocol. Available online: https://academictorrents.com/details/817b91fd639263f6f644de4ccc9575c20b005c6c (accessed on 30 May 2023).
36. Mohanty C, Mahapatra S, Acharya B, et al. Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy. Sensors. 2023, 23(12): 5726. doi: 10.3390/s23125726
37. Baswaraju S, Maheswari VU, Chennam K, et al. Future Food Production Prediction Using AROA Based Hybrid Deep Learning Model in Agri-Sector. Human-Centric Intelligent Systems. 2023, 1-16.