Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 27s (2025)
https://www.jneonatalsurg.com

An Optimal Transfer Learning Approach for Diabetic Retinopathy Severity Level Detection with Cluster Based Image Segmentation

C.Sivaranjani[1], Dr.C. Jeyabharathi[2], Dr.S. Vimala[3]

[1]Ph.D Research Scholar, Mother Teresa Women's University, Kodaikanal, Tamilnadu, India

Email ID: c.s.ranjanii@gmail.com

[2]Computer Instructor Grade - I (PG), Govt. Higher Secondary School, Thalaiyuthu, Palani, Tamilnadu

Email ID: bharathi_guhan@yahoo.com

[3]Associate Professor, Dept. of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamilnadu, India

Email ID: vimalaharini@gmail.com

Cite this paper as: C.Sivaranjani, Dr.C. Jeyabharathi, Dr.S. Vimala, (2025) An Optimal Transfer Learning Approach for Diabetic Retinopathy Severity Level Detection with Cluster Based Image Segmentation, *Journal of Neonatal Surgery*, 14 (27s), 570-580

ABSTRACT

Diabetic Retinopathy is a mutual snag resulting from diabetes, which affects the retina. It is the leading cause of blindness worldwide and detecting this disease early can prevent it from causing patients to lose their sight. But the early detection of Diabetic Retinopathy is very tough and requires an expert interpretation on fundus images by clinical specialists. Conclusions: A novel deep learning-based diabetic retinopathy classification with integrated data balancing transfer learning for the feature selection approach was recommended in this study. The system includes 4 phases including; dataset balancing, preprocessing, segmentation and classification. The system initially carries out data balancing by the Near-miss algorithm. The preprocessing is followed up where the images are enhanced by Retinex algorithm and image augmentation is performed for enhancing the dataset quality. After which, the Deviation cantered K-Means (DKM) algorithm is used to segment. Useful features are extracted and classified from the segmented images using Optimal Vector Pooling Centred Exception Network (OVPXNet). We tested and validated the system to attain higher accuracy of 99.66% for MESSIDOR-2 dataset, average accuracy achieved: i.e., APTOS &MESSIDOR-2 was 99.59%.

Keywords: *Diabetic Retinopathy, Deep Learning, Pre-Trained Model, Dataset balancing, Segmentation, Feature Extraction, and Retinal Fundus Images*

1. INTRODUCTION

Diabetic Retinopathy is a micro vascular problem of diabetes, which currently suffers the obvious risk of becoming one in future cause of vision loss and blindness [1]. This disease is characterized by structural and functional changes in retinal microvasculature, featuring capillary occlusion, leakage as well as retinal ischemia and neovascularization [2]. The International Diabetes Federation estimates that the number of people with diabetes is going to rise from 285.1 million in 2010, and will reach up to about 191.0 million by year 2030 [3]. It is a lingering progressive sickness that matures from mild NPDR to moderate NPDR then into its irreversible form which called as DR, Figure - 02. DP can be broadly classified. Classically, NPDR is graded on the basis of clinically observable pathologic phenotypes such as micro aneurysms (MAs), haemorrhages (HEs), exudates and arteriolar abnormalities [4]. This will help the persons to get proper medication and advice regarding diabetes so that its effect can be minimized by early detection, diagnosis, and treatment [5]. Fundus photography as a well-tolerated, rapid, non-invasive and widely available tool is the most commonly used method to investigate DR coverage [6]; actually manual identification of DR from fundus images requires high degree of expert insight & effort shown by an eye-specialist ophthalmologist. Earlier CAD systems are becoming crucial for automated identification of DR in fundus images [7]. Although low-contrast medical images are popular in CAD systems based on various imaging techniques to improve the overall performance of DR detection [8], it remains challenging to detect differences between retinopathy-related fundus images [9].

Artificial intelligence (AI) is demonstrated to be used as a conduct's suppressor of DR and course on vision loss [10] ML is a sub-section of AI and can be described as the imitation of intellectual human behaviour by machine. These algorithms can provide multidimensional solutions to the outcomes with more accuracy as compared to manual diagnosis [11]. Sustenance path engine, result bush, logistic deterioration and accidental woodland (RF) are the most common ML algorithms for classifying the DR [15], but their performance is poorly practical, they have developed very slowly in recent years. Additionally, the process of feature extracting and selection is time-consuming and subject to different objects. DL has emerged as a breakthrough technology in automatic feature engineering, which is capable of solving the vanishing gradient problem by learning about composite features without having to design them while it also learns incrementally. DL is a type of AI which runs within neural networks having numerous processing layers that extract the features from data in deep level [12]. CNN has become the state-of-the-art DL technique for DR detection [13]. With respect to prediction and detection on the medical image's identification, it has given phenomenal outcomes. Though it runs much faster compared to something like Maxol because of an operation such and if the CNN has multiple layers, then this training process fries a lot and can be made both times consuming as well hardware dependent.

GRU, CNN, RNN, LSTM etc are some other famous DL techniques used to detect DR. These techniques are designed to get the best and accurate prediction without human intervention which results in early predictions of diseases at initial stages making billions lives saved [14]. The diagnostic potential for DR has further improved with a number of studies combining transfer deep learning models, such as VGG [16], Reset [17] Alex Net [18] Inception[19-21]) and Exception etc. These approaches offer a reduction in training time and resource demands, as compared to initializing all from scratch which might make it feasible for his system is running on very limited hardware; however more improvement still needed. Therefore, for efficient DR classification, the proposed system leverages enhanced transfer learning based feature rich representations. The main contributions of the paper are listed as follows:

- The collected MESSIDOR-2 and APTOS datasets are balanced with the Near-Miss algorithm, which views the class distribution and eliminates samples sited in the higher class.
- The contrast of the image is improved by Retime and generates good visual scenes.
- The DKM algorithm is used to segment the important parts in images, optimizing the physician's diagnosis as extensive as perceptive the algorithm spots the suspicious patterns.
- OVPXNet algorithm is used to undertake DR classification by drawing useful features from the segmented images to efficiently reduce the data. Vector pooling is incorporated into existing ZNet, henceforth improving the prediction efficiency and reducing their computational resources. Hyper parameters are significantly chosen with an SBMO algorithm to minimize the classification loss.

The manuscript is divided into multiple parts. Related work section discusses some of the related work on early DR detection proposed prior models, methods or research. Detailed example of planned way is covered in Section 3. In section 4, we appraise the recital of our model with other existing methods and finally in section 5, conclusion closes this paper along with future directions.

2. LITERATURE SURVEY

Many researchers have used retinal fundus photography images on DR detection and diagnosis, additionally the ML & DL method has been used to detect DR by **T. Vijayan et al.**, Feature Selection for Simple Colour Histogram Filter on Retinal Fundus Image of Diabetic Retinopathy [16]. Preprocessing was implemented initially to remove the noise from images. After which again the features were extracted from simple colour histogram filter and then final DR classification is performed based on Decision Tree (DT) & k-nearest neighbour (KNN). The results of the KNN model obtained best accuracy equal to 81.99% comparing with DT method and it is performed on one experiment in retinal Kaggle dataset Ghulam

G. Kalyani et al. [17] proposed DR diagnosis and classification with Capsule networks. Preprocessing - Here Messi or dataset which was collected is pre-processed with CLAHE for improving the contrast of image and Gaussian filter to remove noise from the image. The capsular network was used to get the features from help edict and extract DR type, achieved an accuracy of 97.98%, 97.65% & FFNET (36%), it predicted that output as healthy retina to fist stage fundus images stages2and3(respectively) are, achieving accuracies of upto 98.64;%

Usharani Bhasvara *et al*. [18] automated detection and classification of DR through the improved pooling function in CNN. Firstly, the system utilized an improved artificial bee colony (ABC) algorithm was employed for enhancement of visual contents about lesions. This was followed by the augmentation to prevent overfitting. Next, the features were extracted using improved ResNet50 with pooling function in CNN. In the end, an improved SVM with linear mapping was employed as a classifier for calculating a feature score based on features extracted and determining lesion types. The performance results showed that the system had an accuracy of 98.32% and 95% for APTOS dataset, and Kaggle contest finalists respectively.

Zongyang Gu *et al.* [19] proposed vision transformer and residual attention for an image, to presented a scale of DR severity in fundus images. Sentence1: The system was comprised of two primary modules - a feature extraction block (FEB), and grading prediction block (GPB). Basically, the FEB applied transformer to mask retinal haemorrhage and exudates areas. It then used RESA to capture the varying global context occupied by different object classes in GPB, shown as middle of

Figure 1. These experiments are performed on the DDR and Indriid data-sets, from which it achieves 0.8235% accuracy than other methods.

Muhaimin Azam Khan Riaan *et al.* [20] trained a lightweight robust deep learning model achieved high accuracy in classifying various diabetic retinopathy images. First, the image preprocessing methods processed images as to clean them of artifacts and other noise factors to make their quality better. Then, three types of augmentations such as geometric, photometric and elastic deformation were applied to obtain a fair dataset. Shallow CNN was then presented [39, 40] to identify the graders of DR. The system was evaluated on three separate DR Datasets: APTOS, Messidor2 and Indriid together, with the architecture given in result 1 of Table -3 using best recorded accuracy.

Md. Nahiduzzaman *et al.* [21] introduced identification of DR with CNN as a feature extractor; and ELM based classifier. Initially, the images were pre-processed using CLAHE for better lesion visibility and then normalized followed by reshaping. Then, a CNN model was constructed to learn the most representing features and an ELM algorithm is utilized for grading in of severity levels or stages DR like normal, proliferative diabetic retinopathy (PDR) etc. The performance analysis testing has been carried out on two datasets i.e., Kaggle DR 2015 and APTOS 2019 respectively by following this technique. The system achieved accuracies of 91.78 % and 97.27 % for the two datasets referred to above respectively.

3. PROPOSED METHODOLOGY

Fig. 1 shows the overall workflow of the proposed DR classification using novel deep learning with data balancing integrated transfer learning for feature learning. The workflow mainly uses the following phases: Dataset balancing, Preprocessing, Segmentation, and Classification. It uses the MESSIDOR-2 and APTOS dataset's images as input to classify the severity of the DR. These phases are briefly explained in the following sub-sections.

Dataset balancing

The initial challenge in the DR classification is that we require a benchmark dataset of lot-labeled retinal fundus images. Therefore, this study uses MESSIDOR-2 and APTOS datasets. Conventional class-imbalanced data leads to bad performances of typical classifiers. The classifiers usually try to minimize global error and do not consider data distribution. Therefore, we can achieve high classification accuracy if all instances are classified as harmful. Thus, to address the above two issues - overfitting caused by a class imbalance or a small effect of misclassification errors- this paper balances the dataset using Near Miss, as discussed previously. Near-miss Algorithm: - It is the under-sampling algorithm used for balancing an imbalanced dataset, which is also one of the better ways to balance data.

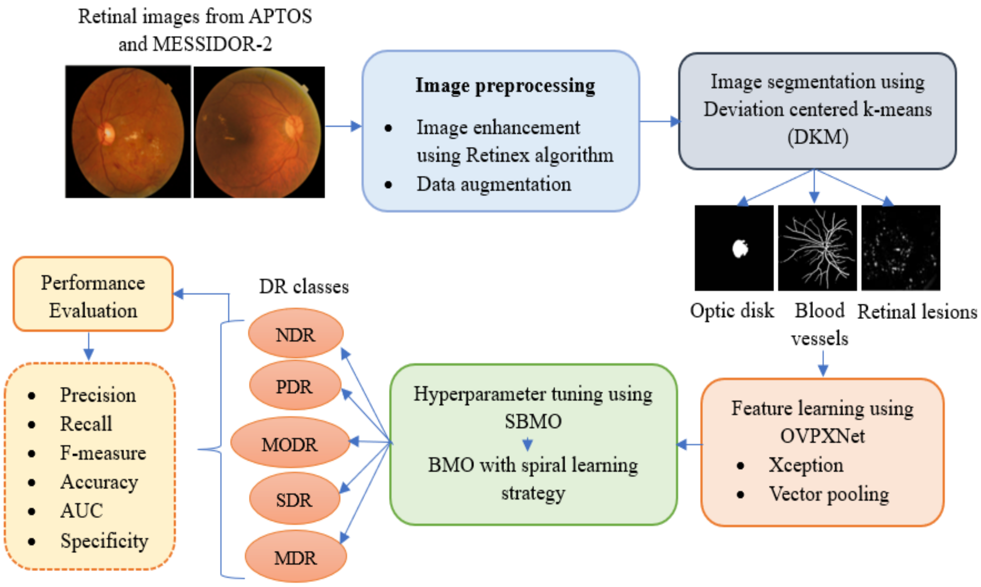

Fig.1. Workflow of the proposed method

The Near-miss algorithm looks at the class distribution and eliminates samples to maintain evenness. In simple terms, if a couple of nearby points (from different classes) should be an outlier, the algorithm removes one from a high class to keep things fair. The stepladders involved in the Near-Miss algorithm are given as below:

- The algorithm first calculates the distance between all the points in the larger class with the points in the smaller class. This can make the process of under-sampling easier.

- Select instances of the larger class that have the shortest distance with the smaller class. These N-classes need to be stored for elimination.
- If there are M- instances of the smaller class then the algorithm will return M x N instances of the larger class.

Preprocessing

These preprocessing methods are used to change the data meaningful. Preprocessing is advantageous for low-contrast illumination, denoising the image testing and their edges sharpening. Preprocessing techniques are applied for the model to perform well. Preprocessing (image enhancement and augmentation) is done on the MESSIDOR-2 and APTOS datasets. These are shortly described as follows:

1) Image Enhancement:

Image enhancement: The initial preprocessing stages were mainly focused on image enhancement. Every retina fundus image underwent these steps via the Retime algorithm. Retime is a nonlinear spectral/spatial transformation that significantly enhances color constancy and local contrast, using various nonlinear illumination conditions to enhance images. These are the parameters you use depending on what kind of images you are using to evaluate [31].This algorithm performs well in promoting the contrasting imagery standard of low-visibility conditions by using the following equation (1).

$$\dot{\epsilon}_{EI}(x, y) = \dot{R}_F(x, y).\, \dot{L}_F(x, y) \qquad (1)$$

Where, $\dot{\epsilon}_{EI}(x, y)$ denotes the contrast enhanced image, $\dot{R}_F(x, y)$ refers the reflectance of the image, and $\dot{L}_F(x, y)$ indicates the illumination of the image, respectively.

*2)**Augmentation**:*

The idea behind using augmentation is to generate more training data and add a bit of distortion in the images which helps prevent overfitting while fitting the model. Every raster of renewed retinal images is processed with the use of image augmentation techniques including rotation, shearing, zooming flipping plus resizing. A rotation-based image augmentation is done by rotating an image around its central plot point. Either side of the axis for rightward or leftward in which its division is [1°,359°]. A shearing operation is movement of one face, in a direction parallel to itself (usually the x or y axis) the result is a parallelogram. A vertical shear moves an edge up or down along the Y-axis, while a horizontal shear slides it left or right on the X-axis. Now we augment the zooming - Zoom Augmentation randomly either zooms in or out of an image. Flipping refers to inverting the image either horizontally or vertically. Finally, the fundus images are resized for normalization reasons. In this resize all the fundus images were converted into [229 × 229].

Segmentation

Segmentation on fundus imaging is an essential step because it contributes to the validated region of interest areas frequently not detectable, which heavily increases the classification performance for DR. To identify the central and local areas of the fovea, optic disk, retinal lesions, or micro-blood vessels in this study, the respective anatomical region is partitioned into several pixels (caused by natural random noise) groups/mosaics. These share some parts in the weight of diabetic retinopathy, significantly enhancing accuracy. However, manual extraction of such features from the fundus images is a time-consuming process and requires considerable investments in terms of skill set. It was compared to other partitional clustering algorithms on some efficiency metrics. However, the method divides a given dataset into k clusters. One complete iteration means when all data objects are assigned to any k clusters, the iterative part is completed, and early grouping is done. It gives higher efficiency when clustering high-dimensional data. However, there are several areas for improvement, such as predefined number and values seeds for clusters, which must be initialized in advance due to random initialization that might converge into local optima. Therefore, the present system uses the deviation theory to choose the optimal cluster centroid. We call the deviation theory the existing KM DKM.

Step 1: Initially, the cluster centroid is selected by using the deviation theory. The deviation measures the spreading of data circulation and the level to which figures ethics contrast from the math mean. It is expressed as charts:

$$\hat{D}_k = \overleftarrow{P}_s - \frac{1}{N}\sum_{k=1}^{N}\overleftarrow{P}_s \qquad (2)$$

Where, $\overleftarrow{P}_s$ indicates a subset in the original pre-processed dataset $\overleftarrow{P}$, $\hat{D}_k$ refers the distance between data in the samples, and denotes the total number of samples, respectively. In this process, the bigger the figures opinion distance, the extra imaginable it will develop a grouping essential idea. We can find the position of the collecting centre more competently

Step 2: After that, analyze the aloofness among each of the figures ideas to each of the hubs, and allot each fact to the contiguous focus.

Step 3: Then, recalculate the newfangled group focus by devious the cruel price of all data ideas in the particular collection.

Step 4: If the cluster centres obtained in step 3 are the same as those obtained in previous iteration, then output the clustering results; otherwise go to step 2.

Detection of DR Grades

We perform the classification of segmented data after segmentation with OVPXNet, where the same network is employed for feature extraction, and then we use it for classification. Mouth Withdrawal: Nose abstraction recognizes and cuttings pertinent topographies since rare information. It makes learning models faster and more accurate. The neural network is trained to classify the dissimilar heights of DR from fundus photographs by means of features extracted. For this purpose, the proposed system has employed a refined Captioned (Net). Exception (or ZNet for short) is a stage model - it stands from extreme inception). These are organized into 14 modules. Every module has a linear residual connection (except the first and last). Input goes through three flows: entry, middle x 8, and exit—batch normalization after each convolution and separable convolutional layer. The middle flow is the central structural component of ZNet and consists of an eight-times-repeated nine-layer structure.

A standard convolutional architecture will be too slow compared to this network, which can extract richer features much faster. One of the critical steps in Net is layer pooling which reduces dimensionality but has a disadvantage that it may discard certain features and partition spatial resolution. To this end, Net employs the vector pooling instead of max pooling and average pooling in [22] to perform dimension reduction for convolution feature maps reliably without losing valuable semantics. Thus, the performance of the net is maximized by minimizing its classification errors and optimized SR hyperparameters via SBMO. Thus, this improvisation included in the existing XNet is terms as OVPXNet and this is shown in Fig 2.

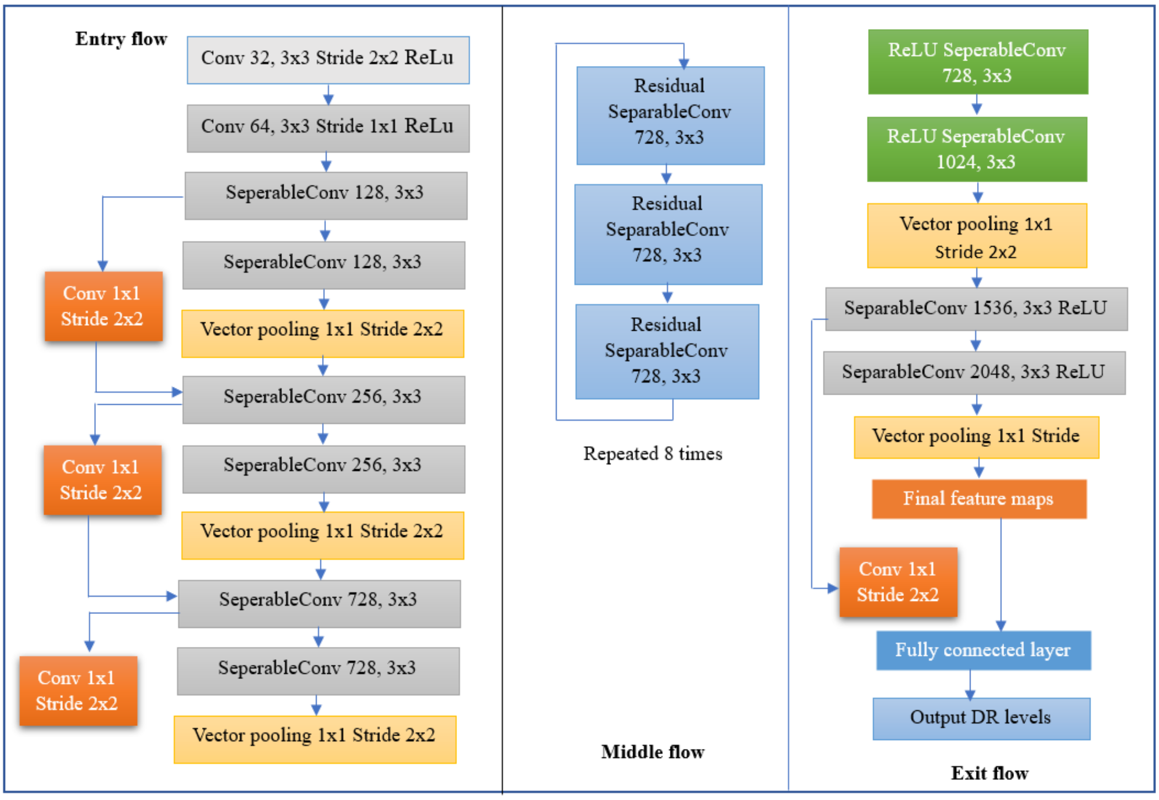

Fig.2. Structure of the proposed OVPXNet

Fig 2 mainly comprises entry flow, middle flow, and exit flow. Exception Entry flow corresponds to the initial part of the model, which is used for the pre-processing of input data and extracting fine-scale features from the input)It is responsible for slowly down sampling the input image and capturing features varying in spatial hierarchies. These segmented images are first passed into the entry flow with a resolution of 299×299×3. In the entry flow, input is first fed into a 1x1 kernel with stride two general convolutional layers. A general convolutional layer is integral to an ZNet and aims to perform feature extraction. It convolves the input image using convolution operators and stores these results into individual channels of a new "convolution" layer. Then comes the separable convolution layer, which takes convoluted features as input. The size of the separable convolutional kernel 3. Grade wide convolutions are separable and happen in depth-wise convolution layers. The depth-wised convolutional means computation and parameters in color channels are carried out through total computation while doing convolutions. The filter goes through a depth-wise convolution to the channels of the input data set and produces its feature map. Batch normalization is used after a depth-wise separable convolutional layer. Batch normalization can decrease overfitting and increase the generalization.

Later, the final output feature maps are sent to the Vector pooling block for dimension reduction of features and keeping essential details about those (size x size) important parts on them. It enters into two equivalent tracks, the Level assembling sheet and the Upright Merging Sheet, respectively. Finally, each pooling path has a 1x1 convolutional sheet with the figure of nuts equal that in the immediate previous layer following the output. That helps to capture dense features is a 1x1 convolutional layer. In the above architecture, each 1x1 density cover in the trajectory sharing hunk is succeeded by REL for more stable performance and faster convergence. Summing or performing a matrix multiplication on the extracted feature vectors of two paths element-wise to keep the pooling layer as dimensionality reduction is one way to combine them. The extracted feature vectors are summed element-by-element to be combined via the vector pooling block to increase efficiency. It can be mathematically expressed as follows:

$$\vec{U}_{vertical} = \xi^{*}\left(Conv_{1\times1}\left(Pool_{1\times N}\left(\vec{I}_{EF}\right)\right)\right) \quad (3)$$

$$\vec{U}_{horizontal} = \xi^{*}\left(Conv_{1\times1}\left(Pool_{1\times N}\left(\vec{I}_{EF}\right)\right)\right) \quad (4)$$

$$\vec{U}_{output} = \xi^{*}\left(\vec{U}_{vertical} \oplus \vec{U}_{horizontal}\right) \quad (5)$$

Where, $\oplus$ refers the element-by-element summation and ξ^{*} REL beginning meaning. The ReLU is a non-linear meaning or piecewise line purpose that will productivity the effort right if it is confident, then, it will production zero by avoiding vanishing gradient issues, ReLU accelerates training convergence. It is formulated as follows:

$$\xi^{*} = \max\left(0, \vec{I}_{EF}\right) \quad (6)$$

The size at the end of the entry flow is 19x19x728 feature maps. After that, the features are sent to the middle flow. The middle flow is occasionally called the core structure part; it consists of a 9-layer in which each layer recurs for further convolution two times, and deeper filters improve the absence of information. Separable convolution and REL activation functions are utilized across the three separable convolutions, all together gaining a study base of deep neural system layers in Table 1. Total 3 × 3 conv-conv layers (728 channels for each layer) in the middle flow, with data tensor repeated eight times. The international open pattern of exports follows the export flow and cooling that comments come after. The flow includes all four separable convolution layers above, flattens the features, and proceeds through this vector pooling block, which is used to reduce dimension, similar to if it was a fully connected layer. The exit flow 3×3 convolution kernel feature number increases from 728 to full-connected final11*11 drop full concert Ch(2048). Features maps are concatenated and fed to a fully connected layer for final classification, which can detect various DR levels from the extracted feature sets. Table 1 shows the layers information of the model.

Table 1: Model parameters used for the proposed VPXNet model

Flows	Layers	Output size	Settings
Entry flow	Conv	299×299×3	Conv 1×1 and stride 2×2
	SeparableConv	256×256×3	Conv 3×3 and stride 2×2
	Vector pooling	19×19×728	Conv 1×1 and stride 2×2
Middle flow	SeparableConv	19×19×728	Conv 3×3 and stride 2×2 (Repeated 8 times)
Exit flow	SeparableConv	19×19×728	Conv 3×3 and stride 2×2
	Vector Pooling	2048-dimen. vector	1×1 and stride 2×2

Hyperparameter tuning

This hyperparameter tuning process will provide us with the best-suited hyperparameters for the Net model. Before the model training, Hyperparameter tuning is done. The hyperparameters selection process is the primary hallmark of BGRU models, which incorporate aspects such as memory and computation complexity. Spiral-based Barnacles Mating Optimization algorithms (SBMO) are used to select the best hyperparameter selection in this network. The concept was inspired by the natural mating behaviour of barnacles, which can occur either through normal copulation or sperm-cast. This, even in two modes, was triggered individually with a BMO. Such as, barnacles have this exclusive behaviour of reproducing new offspring; solving combinatorial optimization is more complex and NP-hard. Its low convergence accuracy and more expensive and time-consuming speed limitation fall into local optima. As a result, the authors have proposed the spiral learning model to address these problems. So, this improvisation in classical BMO is called SBMO. The process involved in the SBMO is explained as follows:

The barnacles are considered as the primary populations for the solution that are generated randomly. After that, compute the fitness of each individual by using equation (13). Minimum classification error is set as the fitness to select the optimal hyperparameters.

$$\begin{aligned} fitness &= Min(classifier\ error) \\ &= Min\left(\frac{No\ of\ misclassified\ samples}{Total\ no\ of\ samples} * 100\right) \end{aligned} \tag{13}$$

Upon close of the restatement point, the process's productivity produces the best key (i.e. optimal hyperparameters) found with its best fitness. The pseudocode of the SBMO is shown in Fig 3.

```
Input: Random set of hyperparameters
Output: Optimal hyperparameters
-----------------------------------------------
Begin
        Initialize the population of barnacles
        Initialize iteration N, maximum iteration (Max_N), and penis length (P_L)
        Compute the fitness of each individual by using Eqn (13)
        While N < Max_N do
                If the selected Dad and Mum ≤ P_L then
                        For each variable do
                                Generate the offspring by Eqn (14)
                        End for
                Else
                        Generate the offspring by Eqn (15)
                End if
                N = N + 1
        End while
End
```

Fig 3: Pseudocode of the proposed SBMO algorithm

4. RESULT AND DISCUSSION

The presentation of the planned scheme for Diabetic Retinopathy classification is examined in some evaluation metrics as compared with hi-tech approaches in this unit. The system has been implemented using C++ 20.0 programming language and Machine Configurations of processor — 8th Group Intel® Core™ i5-8250U CPU with base frequency1.60GHz, Turbo Boost Frequency up to3.40 GHz,6MB Cache Memory – 16GB Storage - 1TB. The hyper parameters used in the proposed work are given in Table 2.

Table 2: Hyperparameters setting

SI. No.	Hyperparameters	Values
1	Optimization	SBMO
2	Activation	ReLU
3	Dropout	0.5
4	Batch size	32
5	Learning rate	0.001

Dataset Descriptions

The presentation of the planned approach was assessed by adopting two public datasets: MESSIDOR and Asia Pacific Tele-Ophthalmology Society (APTOS) datasets. The Messidor-2 dataset is a publically available dataset and a gathering of DR inspections, each entailing of two macula-centred judgment fundus descriptions (one per eye).It has been globally used by researchers for DR detection algorithm analysis, which is an extension of Messidor. It contains 1,748 retinal images of 874 examinations. Although there are no official annotations for this dataset, the third-party grades for 1,744 out of the 1,748 images adjudicated by a panel of three retina specialists are available for researchers. It contains five groups: No Diabetic Retinopathy (NDR), mild DR (MDR), moderate DR (MODR), severe DR (SDR), and proliferative DR (PDR). It is access through https://www.adcis.net/en/third-party/messidor2/.

We have used the APTOS dataset, which contains 3662 publically available annotated retinal images of varying resolution and quality, acquired using various imaging types in multiple participants from rural India. Results: This dataset was curated by the Arvind Eye Hospital, India. The fundus photographs were gathered under many circumstances and conditions over time. Subsequently, the collected samples were examined and graded by trained graders according to the guidelines of the International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRSS). According to the APTOS scaling system, the samples are categorized as seen in Table 1: NDR - No Diabetic Retinopathy MDR - Mild No proliferative Diabetic Retinopathy MODR- Moderate No proliferative DR SDR- Severe no proliferative diabetic retinopathy (as per haggle train and test split) PDAR - Mention NP block DR partially proliferated. It is available on the Internet and it is access through using the following https://www.kaggle.com/datasets/mariaherrerot/aptos2019

Performance Analysis under MESSIDOR-2 Dataset

In this section, the outcomes of the proposed SHBGRU model are investigated against the existing XNet, Deep CNN, RNN and XGBoost methods in terms of the above-mentioned metrics under the MESSIDOR-2 dataset, which is given in the following table and figures.

Table 3: Results analysis based on (a) accuracy, (b) precision, (c) recall, and (d) f1-score

(a)

Techniques	NDR	MDR	MODR	SDR	PDR
Proposed	99.72	99.48	99.53	99.75	99.84
XNet	97.52	97.23	97.35	97.55	97.65
CNN	95.64	95.35	95.45	95.67	95.77
RNN	93.41	93.12	93.26	93.44	93.56
XGBoost	91.36	91.09	91.13	91.39	91.45

(b)

Techniques	NDR	MDR	MODR	SDR	PDR
Proposed	99.81	99.54	99.57	99.84	99.91
XNet	97.61	97.32	97.41	97.62	97.73

CNN	95.72	95.41	95.52	95.74	95.85
RNN	93.49	93.21	93.35	93.51	93.64
XGBoost	91.42	91.18	91.21	91.46	91.52

(c)

Techniques	NDR	MDR	MODR	SDR	PDR
Proposed	99.65	99.39	99.45	99.68	99.76
XNet	97.45	97.18	97.28	97.49	97.59
CNN	95.56	95.28	95.39	95.61	95.71
RNN	93.35	93.08	93.19	93.38	93.49
XGBoost	91.29	91.01	91.07	91.31	91.38

(d)

Techniques	NDR	MDR	MODR	SDR	PDR
Proposed	99.75	99.51	99.57	99.78	99.87
XNet	97.55	97.27	97.38	97.58	97.69
CNN	95.67	95.38	95.49	95.81	95.79
RNN	93.44	93.15	93.29	93.47	93.59
XGBoost	91.39	91.12	91.15	91.42	91.48

Table 3 shows the tremendous outcomes of the proposed model. Classification results are presented for five classes: NDR, MDR, MODR, SDR, and PDR. Table 3 (a): In Table 3(a), For NDR classes, the proposed accuracy is higher than existing methods in comparison to the accuracies we have taken from literature and are as follows:-97.52%, 95.64%, 93.41%, 91.36% (proposed) 99.72% (NDA)-(Most promising result). On the other hand, the suggested model also attains superior classification accuracy over all those left classes, such as MDR, MODR, and SDR & PDR. In addition, in Table 3(b), (c), and (d), the proposed one also gets better results in precision, recall, and f1-score rate compared to existing methods. Fig 4 shows the average results vs. best of gene, parameter, error rate, AUC, and specificity for proposed against existing models. The existing Boost is implemented more slowly in this figure than all the other proposed and existing methods. Another example is the fact that less predictive results, where we achieve 91.28% accuracy (Acc), precision (Pre), and recall (Rec), are offered by it, but the best scores of our proposed model reached up to 99.66%, yield in an f1-score, etc. and Thus, it shows that the proposed one outperformed existing methods.

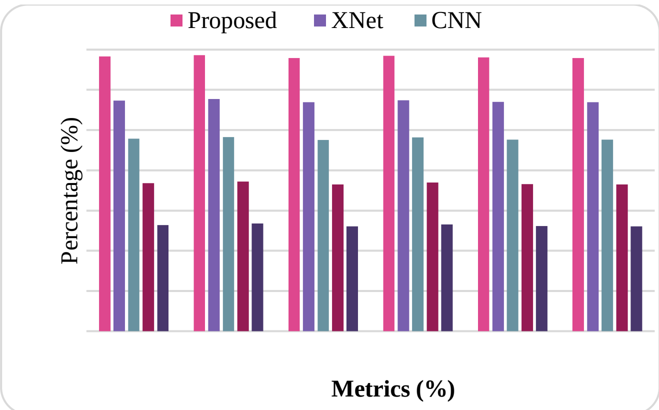

Fig 4: Average results of the proposed model

Performance Analysis under APTOS Dataset

In this section, the outcomes of the proposed SHBGRU model are weighted against the existing methods in terms of the above-mentioned metrics under the APTOS dataset, which is shown in Fig 5. It shows the efficiency of the recommended approach, which is done based on the five classes of DR such as NDR, MDR, MODR, SDR, and PDR

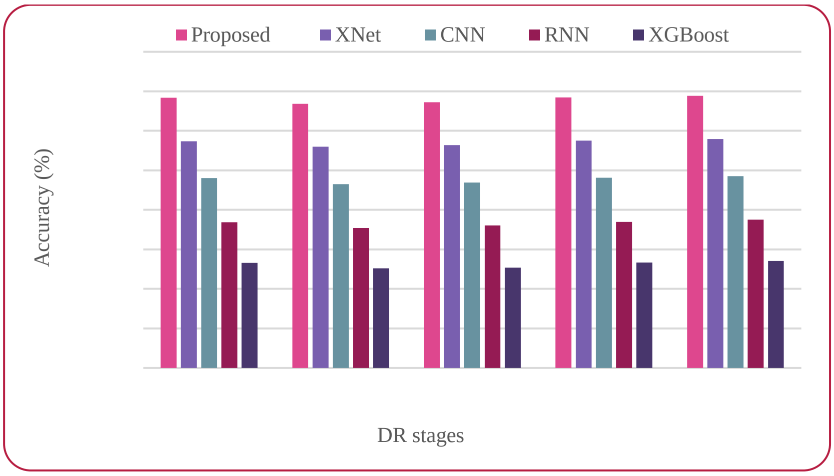

Fig 5: Result analysis

The proposed one achieves an accuracy of 99.68%, 99.36%, 99.45%, 99.69%, and 99.77% for the five classes, a precision of 99.75%, 99.42%, 99.51%, 99.76%, and 99.83% for five classes, recall of 99.61%, 99.28%, 99.38%, 99.61%, and 99.69% for five classes, and f1-score of 99.71%, 99.39%, 99.48%, 99.72%, and 99.81% for five classes, respectively, which is better than the existing methods.

5. CONCLUSION

In this study, we proposed a deep learning approach aiming to classify diabetic retinopathy and present potential solutions further differentiating validation errors from training caveats using data-balancing integrated transfer-learning-based feature learning. It evaluates the system's effectiveness, and these experiments are performed using MESSIODR-2 and APTOS datasets. Moreover, we compare the efficiency of our proposed model with Net CNN RNN Boos methods. The evaluation uses the most common metrics: accuracy, precision, recall f1-score AUC specificity, and error rate. The performance of the proposed model, along with existing ones, is firstly analyzed on the MESSIDOR-2 dataset concerning classification stages such as NDR, MDR, MODR, SDR, and PDP, respectively, summarized where shown in Table 2. The proposed obtained an average result with 99.66% accuracy, 99.73% precision, 99.59% recall, and f1-score rate values of the prediction model applied to detect AS in particular conditions with AUC=0.98. These results translate into a specificity of about 84-88%, which is mentioned as general statistics. Second, the proposed model is compared with existing methods based on whether it can improve their efficiency for the APTOS dataset. The projected classical bring about in an average classification accurateness of 99.59%, a precision value of 99.65 %, recall equal to 99.51%. In aggregate, the held-out results from this planned classical seemed to do better than all of previous models when trained on both datasets. This could detect the DR earlier in samples, thereby minimizing vision loss of patients and saving time for a health care professional

REFERENCES

[1] Athira, T. R., & Nair, J. J. (2023). Diabetic Retinopathy Grading from Color Fundus Images: An Autotuned Deep Learning Approach. *Procedia Computer Science*, *218*, 1055-1066.

[2] Shi, D., Zhang, W., He, S., Chen, Y., Song, F., Liu, S., ... & He, M. (2023). Translation of color fundus photography into fluorescein angiography using deep learning for enhanced diabetic retinopathy screening. *Ophthalmology Science*, *3*(4), 100401.

[3] Shaukat, N., Amin, J., Sharif, M. I., Sharif, M. I., Kadry, S., & Sevcik, L. (2023). Classification and Segmentation of Diabetic Retinopathy: A Systemic Review. *Applied Sciences*, *13*(5), 3108.

[4] Aziz, T., Charoenlarpnopparut, C., &Mahapakulchai, S. (2023). Deep learning-based hemorrhage detection for diabetic retinopathy screening. *Scientific Reports*, *13*(1), 1479.

[5] Kadry, S., Crespo, R. G., Herrera-Viedma, E., Krishnamoorthy, S., & Rajinikanth, V. (2023). Deep and handcrafted feature supported diabetic retinopathy detection: A study. *Procedia Computer Science*, *218*, 2675-

2683.

[6] Palaniswamy, T., &Vellingiri, M. (2023). Internet of Things and Deep Learning Enabled Diabetic Retinopathy Diagnosis Using Retinal Fundus Images. *IEEE Access*, *11*, 27590-27601.

[7] Thanki, R. (2023). A deep neural network and machine learning approach for retinal fundus image classification. *Healthcare Analytics*, *3*, 100140.

[8] Nagpal, D., Alsubaie, N., Soufiene, B. O., Alqahtani, M. S., Abbas, M., &Almohiy, H. M. (2023). Automatic Detection of Diabetic Hypertensive Retinopathy in Fundus Images Using Transfer Learning. *Applied Sciences*, *13*(8), 4695.

[9] Attallah, O. (2023). GabROP: gabor wavelets-based CAD for retinopathy of prematurity diagnosis via convolutional neural networks. *Diagnostics*, *13*(2), 171.

[10] Grzybowski, A., Singhanetr, P., Nanegrungsunk, O., &Ruamviboonsuk, P. (2023). Artificial intelligence for diabetic retinopathy screening using color retinal photographs: from development to deployment. *Ophthalmology and Therapy*, *12*(3), 1419-1437.

[11] Sangeetha, K., Valarmathi, K., Kalaichelvi, T., &Subburaj, S. (2023). A broad study of machine learning and deep learning techniques for diabetic retinopathy based on feature extraction, detection and classification. *Measurement: Sensors*, *30*, 100951.

[12] Sebastian, A., Elharrouss, O., Al-Maadeed, S., &Almaadeed, N. (2023). A Survey on Deep-Learning-Based Diabetic Retinopathy Classification. *Diagnostics*, *13*(3), 345.

[13] Sudha, S., Srinivasan, A., & Devi, T. G. (2023). Cross-Validation Convolution Neural Network-Based Algorithm for Automated Detection of Diabetic Retinopathy. *Computer Systems Science & Engineering*, *45*(2).

[14] Uppamma, P., & Bhattacharya, S. (2023). Deep Learning and Medical Image Processing Techniques for Diabetic Retinopathy: A Survey of Applications, Challenges, and Future Trends. *Journal of Healthcare Engineering*, *2023*.

[15] Shoaib, M. R., Emara, H. M., Zhao, J., El-Shafai, W., Soliman, N. F., Mubarak, A. S., ... &Esmaiel, H. (2024). Deep learning innovations in diagnosing diabetic retinopathy: The potential of transfer learning and the DiaCNN model. *Computers in Biology and Medicine*, *169*, 107834.

[16] Vijayan, T., Sangeetha, M., Kumaravel, A., & Karthik, B. (2023). Feature selection for simple color histogram filter based on retinal fundus images for diabetic retinopathy recognition. *IETE Journal of Research*, *69*(2), 987-994.

[17] Kalyani, G., Janakiramaiah, B., Karuna, A., & Prasad, L. N. (2023). Diabetic retinopathy detection and classification using capsule networks. Complex & Intelligent Systems, 9(3), 2651-2664.

[18] Bhimavarapu, U., Chintalapudi, N., &Battineni, G. (2023). Automatic Detection and Classification of Diabetic Retinopathy Using the Improved Pooling Function in the Convolution Neural Network. *Diagnostics*, *13*(15), 2606.

[19] Gu, Z., Li, Y., Wang, Z., Kan, J., Shu, J., & Wang, Q. (2023). Classification of diabetic retinopathy severity in fundus images using the vision transformer and residual attention. *Computational Intelligence and Neuroscience*, *2023*.

[20] Raiaan, M. A. K., Fatema, K., Khan, I. U., Azam, S., ur Rashid, M. R., Mukta, M. S. H., ... & De Boer, F. (2023). A Lightweight Robust Deep Learning Model Gained High Accuracy in Classifying a Wide Range of Diabetic Retinopathy Images. *IEEE Access*.

[21] Nahiduzzaman, M., Islam, M. R., Goni, M. O. F., Anower, M. S., Ahsan, M., Haider, J., & Kowalski, M. (2023). Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier. *Expert Systems with Applications*, *217*, 119557.