Received: September 12, 2024. Revised: October 2, 2024.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Classification of Diabetic Retinopathy Stages Using Dual Attention Mechanism with EfficientNet-B0

Niranjan Kakkepalya Ramakrishnaiah[1]* **Nagesh Rangahanumaiah[2]** **Vani Anantharamaiah[1]**
Beena Ullala Mata Bangalore Nanjundaswamy[3] **Krishna Murthy Kodagihalli Timmappa[4]**

[1]*Department of Medical Electronics Engineering, B.M.S. College of Engineering, Bengaluru, India*
[2]*Department of Electronics and Communication Engineering, Govt. SKSJT Institute, Bengaluru, India*
[3]*Department of Electronics and Communication Engineering, B.M.S. College of Engineering, Bengaluru, India*
[4]*Department of Electronics and Instrumentation Engineering, B.M.S. College of Engineering, Bengaluru, India*
* Corresponding author's Email: niranjana.kr@gmail.com

Abstract: Currently, Diabetic Retinopathy (DR) is one of the most common causes of blindness among people who suffer from diabetes for a long period. But, the existing detection models have drawbacks in classifying different stages of DR due to the similarity between some classes. To overcome this problem, a Dual Attention Mechanism with EfficientNet-B0 (DAM-EfficientNet-B0) is proposed for precise detection and classification of DR. Initially, retinal images are obtained from EyePACS and Messidor datasets and fed for preprocessing by Contrast Limited Adaptive Histogram Equalization (CLAHE) and data augmentation. After that, the preprocessed images are segmented based on a modified expectation maximization technique, while the segmented images are fed to the feature extraction process by ResNet-50. Finally, the extracted features are forwarded to the proposed classifier approach for accurate detection and classification of DR as per severity levels. The utilization of DAM helps select the relevant and important features which further leads to learning differences between classes. The experimental results show that DAM-EfficientNet-B0 attains an accuracy of 99.21% which is greater than the previous methods such as MultiStream-Deep Neural Network (MS-DNN), Concatenated Convolutional Neural Network (CCNN), and DNN-Butterfly Optimization Algorithm (DNN-BOA).

Keywords: Contrast limited adaptive histogram equalization, Data augmentation, Diabetic retinopathy, Dual attention mechanism with efficientNet-B0, Modified expectation maximization, ResNet-50.

1. Introduction

Diabetic Retinopathy is an eye disorder that occurs due to the break of retinal blood vessels in the human eye, caused by diabetic mellitus. The DR is generally observed in patients who have diabetes for a longer period (10–15 years), due to the formation of several irregular retinal lesions in the eye [1, 2]. The duration of diabetes of a person is a main factor for DR, and when duration increases, the risk of DR also increases [3]. This is a damage or break of blood vessels in retinal tissue, which is the photosensitive part of our visual apparatus [4]. The oxygen carried to the retina tissues is decreased due to the reduced size of blood vessels, which are ruptured by the weakness of new vessels [5]. The various symptoms of DR are aneurysms, abnormal growth of new blood vessels, nerve tissue damage, blood vessel leakage, and retinal swelling [6]. Patients with diabetes are generally unaware of the possibility of DR, resulting in delayed diagnosis and treatment [7, 8]. Early detection and accurate classification of DR help in preventing people from loss of vision and reduce the severity of DR.

Deep Learning (DL) is the most widely used approach in many applications, exclusively in image processing and classification [9]. In recent times, DL methods have attained great success in the area of computer vision by modelling high-level intellections in data relative to the exact prediction tasks [10]. Effective results have been achieved from the existing DL methods in terms of direct detection and grading of DR associated with traditional machine learning approaches [11]. Nonetheless, DL architectures utilized for DR detection and grading studies have achieved successful results, but with their own limitations [12]. Some of the limitations of the existing approaches include insufficient data for each stage of DR, poor quality of images, and irrelevant features that impact the detection of DR [13-15]. Hence, the existing models struggle to distinguish between severity levels of DR that results in misclassification. To overcome these problems, a DAM- EfficientNet-B0 is proposed for precise detection and classification of retinopathy stages. By utilizing the dual attention mechanism, the irrelevant features are eliminated and more precise information about severity level is obtained to avoid misclassification.

The main contributions of this research are:

- For enhancing image quality CLAHE technique is used to improve contrast and then segmenting retinal images, by a modified expectation maximization is employed to segment exudates that can be identified. An adaptive cluster technique is utilized to identify and remove the fluid-like substance exudates effectively.
- ResNet-50 is employed in the feature extraction process which adds a 3-layer bottleneck block for every two layers that have a better representational capacity to extract the most significant features.
- DAM-EfficientNet-B0 is proposed for DR detection and classification which efficiently detects and distinguishes the retinopathy severity levels effectively. The integration of two attention mechanisms eliminates irrelevant features and focuses only on the important features which leads to precise classification of severity levels.

The research paper is structured as follows: Section 2 explains the literature review, Section 3 describes the methodologies implemented for this research, while the experimental results are illustrated in Section 4, and the conclusion of this research is given in Section 5.

2. Literature review

In this section, the advantages and limitations of the existing detection and classification approaches utilized for diabetic retinopathy are discussed.

Mustafa [16] developed a multi-stream model for severity classification in DR by a boosting framework. The developed model utilized ResNet-50 and DenseNet-121 neural network models for automatic grading and classification of DR. The utilization of ResNet & DenseNet-121 improved the feature extraction process by identity mapping and feature reuse to dense layers which increased the performance accuracy. However, the inadequacy of data in some classes led to the degradation of the detection and classification accuracy of the DenseNet-121 model.

da Rocha [17] designed a classification model for DR based on VGG19 network. The designed DL model classified the preprocessed images into 5 categories with an additional class for reporting low-quality retinal images. The data balance in preprocessing provided adequate data that associated with hyper parameters improved the classification accuracy of the model. There are various parts present in the eye, and especially exudates in the retinal images which are not segmented in the designed model, hence offering inaccurate classification.

Syed and Durai [18] explored a DL based diagnosis model for DR detection and classification. The explored model utilized CCNN for training and LR for multi-classification of DR. An optic disc and blood vessel segmentation in this model helped in the detection and classification of all retinopathy stages. However, the explored CCNN model was sensitive to image quality and struggled to distinguish between different levels of DR which resulted in inaccurate classification.

Kalyani [19] presented a model for the detection and classification of DR based on capsule networks. In the Capsule Network, the initial two layers were utilized for feature extraction and the class capsule layer for the detection of the probability of a specific category. An advantage of the presented Capsule Network model was that it utilized an accurate early detection for better diagnosis which was performed by CapsNet. Nonetheless, the capsule network model had the drawback of classifying other retinopathy stages due to the similarity between severity levels of DR.

Rachupudi [20] introduced an optimized DL model for the detection of DR with retinal fundus images. For automatic diagnosis of DR, DNN with BOA optimizer was utilized to classify the different retinopathy stages with relevant features extracted by grey level co-matrix method. The main advantage of the introduced DNN with BOA model was that the blood vessels and exudates were removed in preprocessing which like fundus, enhanced classification performance.

However, the extracted features with irrelevant information led to inaccurate classification that increased the delay in the early diagnosis of DR.

P. Saranya [21] represented a CNN model for detection and classification of red lesions in retinal images to detect DR. The represented CNN based DR detection model utilized a U-Net model for segmentations of red lesions effectively. The main advantage of the represented CNN model extracts the significant features through the pooling layers from semantic segmented images for detection of DR. however, the U-Net segmentation method focused only on segmenting red lesions not the microaneurysms and haemorrhages which affect the in precise DR detection.

There are some limitations in the existing works are noted as follows: insufficient data for certain classes, retinal images with various parts not segmented, struggles to differentiate between retinopathy stages, and inappropriate features leading to misclassification. In order to overcome these limitations, a DAM-EfficientNet-B0 is proposed for the precise detection and classification of DR. The image quality is enhanced by CLAHE and balance data for all classes to improve the effective detection of DR.

3. Classification of DR stages using DAM-EfficientNet-B0

The proposed detection and classification of DR includes 5 stages: Dataset, Preprocessing, Segmentation, Feature Extraction and Classification. The block diagram of the proposed model is illustrated in Fig. 1. Initially, retinal images are acquired from EyePacs and Messidor datasets and preprocessed by CLAHE to enhance contrast of the image. Then, the enhanced image is segmented by the removal of background and then the feature is extracted. Finally, retinopathy severity stages are classified by the proposed DAM-EfficientNet-B0.

3.1 Datasets

In this research, benchmark datasets: EyePACS and Messidor are utilized for the precise detection of DR.

3.1.1. EyePACS dataset

EyePACS is the most widely utilized dataset for DR detection available on the Kaggle website [22]. This dataset has 88,702 retinal images collected from primary care clinics in California and various regions of the world. The dataset is further categorized into 35,126 images for the training set and 53,576 images for the testing set. These images are classified according to the DR stages of Normal, Mild, Moderate, Severe, and Proliferative, which is shown in Fig. 2.

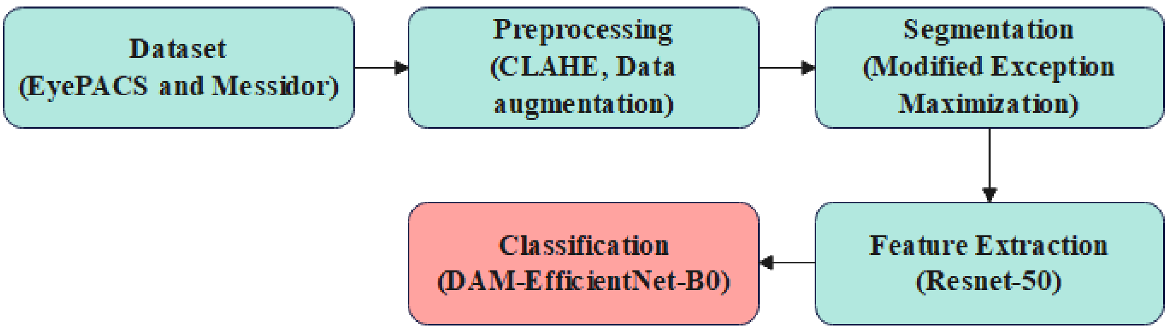

Figure. 1 Block diagram for the proposed model

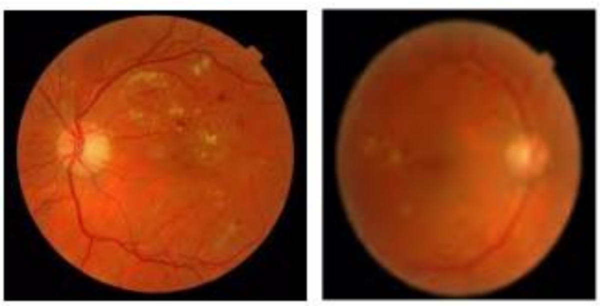

Figure. 2 Retinal images of EyePACS dataset

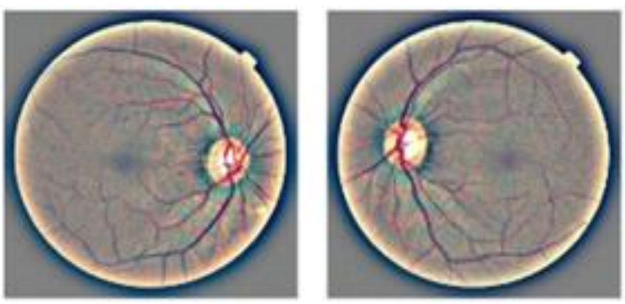

Figure. 3 Retinal images of Messidor dataset

Received: September 12, 2024. Revised: October 2, 2024.

3.1.2. Messidor dataset

Messidor dataset includes 1200 RGB fundus images assimilated by 3 ophthalmologic departments [23]. These images are acquired with 8 bits per color, having the resolution of 1440×960, 2240×1488, or 2304×1536 pixels. In this dataset, 800 images are acquired with pupil dilation, and the remaining 400 images without dilation. The dataset consists of four classes 0, 1, 2, and 3, where 0 denotes normal or no DR, and the remaining three classes indicate severity levels with 1 denotes minimum, and 3 denotes maximum. These images are fed to preprocessing stage to improve raw data for DR detection. The 4 classes of DR in Messidor dataset are represented in Fig. 3.

3.2 Preprocessing

The retinal images obtained from above-mentioned datasets are fed as input to the preprocessing phase to enhance image quality as well as balance the data of all classes for accurate classification. In this research, CLAHE is used to improve contrast in retinal images while enhancing image quality and boosting the detection and classification performance for different severity levels [24]. At first, the retinal images are divided into small parts by nominal separation and then histogram equalization is applied to each separation. By equalizing the grey value in the image, hidden features of the shape of the image are more visible. The sharp edges of the image are protected by boundaries in an area and a grey spectrum is used to illustrate the image. Here, CLAHE is utilized to limit amplification that is determined using a clipping histogram at a certain level. Rayleigh distribution is utilized for clipping histogram which is mathematically expressed in Eq. (1).

$$g = g_{min}\left[2(\beta^2)ln + \left(\frac{1}{1-p(F)}\right)\right]^{0.5} \quad (1)$$

Where, g and g_{min} indicate the measured and minimum pixel values, β represents threshold parameter (clip parameter), and $p(F)$ indicates distribution function. These enhanced retinal images are fed to the next phase of data augmentation that balance data in both datasets.

3.2.1. Data augmentation

The enhanced images are forwarded to the data augmentation process to solve the data imbalance problem in both DR datasets. The data augmentation techniques utilized for this research are: shearing, rotating, cropping, translation process and flipping [25]. The final augmented images are further passed to the segmentation process.

3.3 Segmentation

The augmented images are given as input to the segmentation process to divide the retinal images into multiple regions for accurate detection. Due to similarity between blood vessels and exudates leads to inaccurate classification of DR where both look identical. A grey value threshold technique is utilized to remove blood vessels that segment according to the possibility that the image has a bimodal histogram. For segmenting retinal images, a modified expectation maximization is employed to segment exudates that are identified in retinal images [26]. Exudates are fluid-like structure which is made of cells, proteins and solid materials that leak out from blood vessels to the neighbouring tissues. After segmenting blood vessels and optic disc, the exudates are further identified to remove from the retinal images. The exudate segmentation by modified expectation maximization technique, as expressed in Eqs. (2) and (3).

$$\left[\alpha_{i,j}^{(t+1)}\right]^2 = \sum_{n=1}^{N} \zeta_{i,j,n}\left(y_{i,n} - x_i^{(t)}\right)^2, \zeta_{i,j,n} \epsilon [0,1] \quad (2)$$

$$\sum_{n=1}^{N} \zeta_{i,j,n} = 1 \quad (3)$$

Where, $y_{i,n} - x_i^{(t)}$ represents finite Gaussian mixture model, and $\alpha_{i,j}$ denotes gross error mode probability. The segmented images are further fed to feature extraction phase for extracting relevant features to classify DR.

3.4 Feature extraction

The segmented images are fed as input to the feature extraction model which is based on the neural network architecture to extract important features. ResNet-50 is a pre-trained neural network used for extracting relevant features from the segmented images that help distinguish between classes. Here, ResNet50 is used for feature extraction which is obtained from the ResNet-34 model and modified by replacing the 3-layer bottleneck block for every two layers in the network [27]. This pre-trained network model is reconfigured from model of Resnet50 where fully connected layers are removed.

At last, a global average pooling layer and a dropout of 0.5 with 2084 dense neurons are added to the network for effective deep feature extraction.

Received: September 12, 2024. Revised: October 2, 2024.

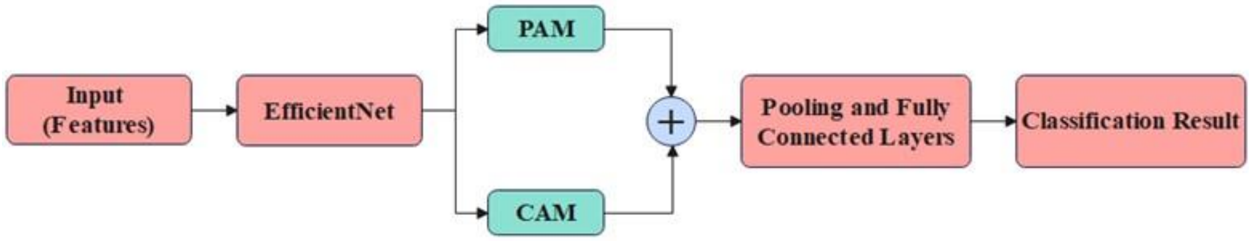

Figure. 4 Structure of DAM-EfficientNet-B0

These extracted features are further passed to proposed classification model for DR detection.

3.5 Proposed classification

The extracted features are taken as input for the classification of DR. For accurate classification, a DAM-EfficientNet-B0 is utilized to learn contextual information by a dual attention mechanism that leads to identifying differences between the retinopathy classes. The main objective of the attention mechanism is to select an important area of input data that does not consider the areas with weak and no discriminative features. These mechanisms focus on distinguishing the features extracted from various regions of input. In this research, DAM is utilized to remove the interference of irrelevant information and focus more on the useful information. Position Attention Mechanism (PAM) and Channel Attention Mechanism (CAM) are incorporated with EfficientNet-B0 to adjust the relationship between local and global features. The structure of DAM-EfficientNet-B0 is illustrated in Fig. 4.

3.5.1. Position attention mechanism

For extracting contextual data from low-level features, PAM is utilized to learn more information by enhancing the representation of features. The PAM structure is designed for collecting and detecting relevant features related to spatial domain. The main advantage of PAM is learning spatial interdependencies of features and obtaining various contextual information from local elements. The input feature map A is expressed as in Eq. (4).

$$A = F\epsilon R^{C\times H\times W} \quad (4)$$

Where, F indicates features; C represents number of relevant features in spatial domain, H denotes spatial dimension height; W indicates the spatial dimension width of the input tensor. The PAM constructs a positional relationship model among features to acquire global feature information and combines features randomly by total weights assigned for features at each position. The input feature map $F\epsilon R^{C\times H\times W}$ executes a convolution operation with a batch normalization layer and a Rectified Linear Unit (ReLU) layer to provide three new feature maps F1, F2, and F3. These maps are called single-channel feature maps which are obtained from F, where feature maps F1 and F2 have the same dimensions, but F3 has different dimensions. After execution of the reshape operation on F1, F2, and F3, the scale of feature maps becomes $H \times W$, then a matrix multiplication is performed. Hence, the final output $F\epsilon R^{C\times H\times W}$ is acquired through the multiplication of distinguished results obtained by learnable parameter δ and an element-wise summation with actual input feature map is given in Eq. (5).

$$F'' = \delta.reshape\left(V.M_p\right) + F' \quad (5)$$

3.5.2. Channel attention mechanism

Channel attention is employed to upgrade features of several channels by simulating the importance of all channels from the extracted feature. The CAM detects relations on the local cross-channel by evaluating the channel and its neighbours. Hence, the number of parameters is minimized and the model complexity is also reduced without compromising the accuracy. In CAM, dense block and transition layers are integrated to avoid adding numerous parameters that lead to overfitting. Transition layer in CAM comprises 1×1 a convolutional layer and an average pooling layer with a stride of 2 for minimizing feature map size. This mechanism is split into 2 phases squeeze and excitation.

In squeeze phase, input features are flattened into one-dimensional vector lengths which denotes number of channels. In the excitation stage, dependencies among channels are collected by a gate mechanism that contains two nonlinear fully connected layers. By squeezing and expanding feature channels, CAM assigns adaptive weights to

 DOI: 10.22266/ijies2024.1231.73

various features. The mathematical representation of the output of CAM is represented in the Eq. (6).

$$\beta(U) = \sigma\left(W_1 W_2\left(U_{avg} + U_{max}\right)\right) \tag{6}$$

Where, W_1 and W_2 are weights of fully connected layer, σ represents the sigmoid activation function, U_{avg} and U_{max} represents average and maximum pooling feature map.

EfficientNet-B0 is a CNN model which utilizes a complex scaling method for image detection and classification. The network is scaled alongside optimizing the depth, width, and resolution that lead to boosting the network. The main objective of EfficientNet-B0 architecture is to find a suitable technique to scale CNN models to achieve better accuracy for the detection recognition or classification process. The authors suggest a compound scaling technique that uses a fixed set of coefficients to scale width, resolution, and depth consistently. When the number of layers (depth) increases, the CNN extract more complex and richer features, but is a challenging task because of the vanishing gradient issue. Thus, scaling width of the network layers, acquiring more fine-grained features and training the model is easier. However, shallow and wide layers in the network are not able to learn from high-level features, and thus high-resolution images are used to learn fine-grained and complex patterns. Hence, adjusting these depth, width and resolution EfficientNet-B0 improves the classification performance. Many convolutional (Conv) layers with 3×3 receptive fields and mobile inverted bottleneck Conv layers acquire information from the extracted features efficiently for DR detection. Scaling depth, width, and resolution in EfficientNet-B0 are represented in Eqs. (7) to (11).

$$d = \alpha\varphi \tag{7}$$

$$w = \beta\varphi \tag{8}$$

$$r = \gamma\varphi \tag{9}$$

$$s.t\alpha.\beta 2.\gamma 2 \approx 2, \tag{10}$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1 \tag{11}$$

Where, $w, d,$ and r represents network's width, depth, and resolutions; constant terms α, β and γ are obtained using the grid search hyperparameter tuning method. A coefficient is a user-defined variable that achieves scaling resources for every model in the network. By fine-tuning the width, depth, and resolution of the network optimize the network accuracy and memory consumption corresponding to the available resources. Unlike deep CNNs, a predefined set of scaling coefficients is used in EfficientNet-B0 to adjust each dimension of layers in the network, which outperforms the previous cutting-edge models trained on ImageNet dataset. An integration of DAM in Efficient-B0 enhances the learning most significant information about severity levels of DR from both datasets and enhances the identification of retinopathy classes effectively. The detected and classified output is obtained from softmax layer of EfficientNet-B0 network accurately. The performance of proposed DR detection and classification model is evaluated by various performance metrics in the section below.

4. Results and discussion

The experimental results and discussion of DAM-EfficientNet-B0 are represented in this section. The system configuration utilized for the simulation of the proposed model is: GPU-4×NVIDIA Tesla M60 (24 cores, 224 GB RAM). The performance metrics employed for the experimental evaluation are: Accuracy, Specificity, Sensitivity, Precision, and F1-score. The mathematical representation of performance metrics is given in Eqs. (12) to (16).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100 \tag{12}$$

$$Specificity = \frac{TP}{TP+FP} \tag{13}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{14}$$

$$Precision = \frac{TP}{TP+FP} \tag{15}$$

$$F1 - Score = \frac{2\times TP}{2\times TP+FP+FN} \times 100 \tag{16}$$

Where, TN is True Negative, FN is False Negative, TP is True Positive, and FP is False Positive.

4.1 Qualitative and quantitative analysis

The qualitative and quantitative analysis of the proposed model is represented from Tables 1 to 6. Table 1 represents the performance of the proposed DAM-EfficientNet-B0 classifier utilizing EyePACS dataset which is estimated by different performance measures. The performance of previous classifier models: CNN, ResNet-50, DenseNet-201, and DAM-EfficientNet-B0 are used for analysis with the proposed method. The DAM incorporated with Efficient-B0 increases the learning of the most significant information about severity levels of DR from both datasets and enhances the detection and classification of retinopathy classes accurately.

Table 2 represents the performance of the proposed DAM-EfficientNet-B0 classifier utilizing Messidor dataset which is estimated by different performance metrics. The performance of previous classifier models: CNN, ResNet-50, DenseNet-201, and EfficientNet are utilized for comparison with the proposed DAM-EfficientNet-B0 model.

The performance of DAM-EfficientNet-B0 classifier is distinguished with different classifier models for DR detection and classification in Table 2. The presented DAM-EfficientNet-B0 classifier attains a superior performance than existing classifier models with accuracy 99.21% on EyePACS and accuracy 99.47% on Messidor datasets. The performance results show that the developed model offers a high classification accuracy than that of the existing DR methods.

Table 3 represents the performance of the feature extraction method utilized for EyePACS dataset which is estimated by different performance measures. The performance of previous feature extraction models namely, ResNet-34, DenseNet-121, and DenseNet-201 are used for comparison with the ResNet-50 method.

Table 4 represents the performance of the feature extraction method utilized for Messidor dataset which is evaluated by different performance metrics. The performance of previous feature extraction models: ResNet-34, DenseNet-121, and DenseNet-201 are used for comparison with the ResNet-50 method.

Table 1. Performance analysis of the proposed method for EyePACS Dataset

Methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**Precision (%)**	**F1-score (%)**
CNN	96.77	95.33	95.59	95.68	95.63
ResNet-50	97.83	96.46	96.54	96.48	96.50
DenseNet-201	98.39	97.52	97.38	97.64	97.50
EfficientNet	98.54	97.79	97.83	97.88	97.85
Proposed DAM-EfficientNet-B0	99.21	98.81	98.74	98.92	98.82

Table 2. Performance analysis of the proposed method for Messidor Dataset

Methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**Precision (%)**	**F1-score (%)**
CNN	96.89	95.36	95.51	95.67	95.58
ResNet-50	97.94	96.29	96.47	96.76	96.61
DenseNet-201	98.43	97.42	97.36	97.59	97.47
EfficientNet	98.68	97.86	97.89	97.92	97.90
Proposed DAM-EfficientNet-B0	99.47	98.92	98.91	98.93	98.91

Table 3. Performance analysis of the Feature Extraction method for EyePACS Dataset

Methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**Precision (%)**	**F1-score (%)**
ResNet-34	96.28	95.74	97.81	97.85	97.82
DenseNet-121	97.46	96.91	96.77	96.82	96.79
DenseNet-201	98.57	97.66	97.51	97.42	97.46
ResNet-50	99.21	98.81	98.74	98.92	98.82

Table 4. Performance analysis of the Feature Extraction method for Messidor Dataset

Methods	**Accuracy (%)**	**Specificity (%)**	**Sensitivity (%)**	**Precision (%)**	**F1-score (%)**
ResNet-34	95.76	94.92	94.91	94.96	94.93
DenseNet-121	97.49	96.63	96.42	96.54	96.47
DenseNet-201	98.58	97.76	97.79	97.82	97.80
ResNet-50	99.47	98.92	98.91	98.93	98.91

Table 5. K-fold analysis of the proposed method for EyePACS Dataset

K-fold	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)	F1-score (%)
EyePACS Dataset					
K=3	96.47	95.79	95.77	95.78	95.77
K=5	99.21	98.81	98.74	98.92	98.82
K=7	97.38	96.59	96.58	96.61	96.59
K=9	98.61	97.83	97.76	97.71	97.73
Messidor Dataset					
K=3	96.38	95.84	95.81	95.86	95.86
K=5	99.47	98.92	98.91	98.93	98.91
K=7	97.82	96.97	96.71	97.68	97.19
K=9	98.94	97.64	97.49	97.38	97.34

Table 6. Comparative analysis of the proposed method for EyePACS and Messidor Dataset

Dataset	Methods	Accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)	F1-score (%)
EyePACS	VGG-16 [17]	77.9	79.1	78.2	79.4	78.6
	CCNN [18]	98.81	N/A	N/A	96.89	89.58
	Proposed DAM-EfficientNet-B0	99.21	98.81	98.74	98.92	99.19
Messidor	CapsNet [19]	97.98	N/A	N/A	95.62	96.36
	DNN-BOA [20]	98.9	98.7	98.3	N/A	N/A
	Proposed DAM-EfficientNet-B0	99.47	98.92	98.91	98.93	98.91
DDR	Multi-stream ResNet and DenseNet-121 [16]	68.24	N/A	N/A	N/A	N/A
	Proposed DAM-EfficientNet-B0	98.46	97.85	96.91	97.92	97.41
IDRiD	CNN [21]	95.65	99	89	N/A	N/A
	Proposed DAM-EfficientNet-B0	98.98	96.44	96.78	97.59	97.18
DIARETB1	CNN [21]	95.5	95.5	95.4	N/A	N/A
	Proposed DAM-EfficientNet-B0	97.77	96.80	96.47	97.21	96.83

Table 5 display the classification results with various K-fold values of the proposed method using different performance measures on EyePACS and Messidor datasets. The dual attention mechanisms with EfficientNet-B0 exhibits commendable outcomes when K-fold value is 5 than when compared to values 3, 7, and 9.

4.2 Comparative analysis

The comparative analysis of the DAM-EfficientNet-B0 method is represented in Table 6. The performance of DAM-EfficientNet-B0 is evaluated with previous DR approaches namely, VGG-16 [17], and CNN [18] in EyePACS dataset. For comparative analysis, four different performance metrics are used on the EyePACS dataset.

The comparative analysis of the proposed method with previous methods confirms that the presented model scores an accuracy of 99.21% and 99.47%, for EyePACS and Messidor datasets which is greater than the previous DL models. The proposed DAM-EfficientNet-B0 is employed for DR detection which accurately detects and classifies the various severity levels in EyePACS dataset effectively. The performance of EfficientNet is evaluated with previous DR detection approaches: CapsNet [19] and DNN-BOA [20] in Messidor dataset. For comparative analysis, four different performance metrics are used for Messidor dataset. The integration of DAM in Efficient-B0 enhances the learning of important information about severity levels of DR and eliminates irrelevant features that improve the identification of retinopathy classes effectively.

4.3 Discussion

The proposed approach accomplishes commendable outcomes with the precise detection and classification of DR on EyePACS and Messidor datasets. DR detection performed by the existing approaches of various ML and DL models has a problem in classifying the DR severity levels due to inadequate data and lack of features which is accomplished by the presented DAM-EfficientNet-B0 method. The VGG-16 [17] data augmentation technique used in this model selects data randomly for classes with minimum data, causing to inaccurate detection. CCNN [18] is sensitive to the quality of the model which impacts the inaccurate detection and classification of DR. CapsNet [19] model fails to distinguish between 4 retinopathy stages. DNN-BOA [20] model struggles to classify emotions accurately because of the irrelevant features. To overcome these limitations, a DAM-EfficientNet-B0 approach is proposed for effective detection and classification of retinopathy stages. With the selected informative features extracted by GLCM, the classification of retinopathy severity classes from EyePACS and Messidor datasets. The dual attention mechanism PAM and CAM alleviates the interference by irrelevant features in EfficientNet-B0 that improve the detection of DR and differentiate the class effectively.

5. Conclusion

The DAM-EfficientNet-B0 is proposed for the detection and classification of DR effectively. Dual attention mechanisms are integrated into EfficientNet-B0 for learning significant information from the extracted features and to remove the interference for improving DR detection. The Efficient-B0 effectively learn from features to identify healthy and diseased retinal images that classify precisely. Then, CLAHE is employed to enhance the contrast of images to distinguish healthy and diseased retinal images for accurate classification. Modified expectation maximization technique is utilized to eliminate exudates which enhances classification by differentiating images as per retinopathy classes. The feature extraction process improved by utilizing ResNet-50 network includes three-layer bottleneck block blocks to learn complex patterns. Finally, the extracted features are fed to proposed classifier model for accurate detection and effective classification of DR by PAM and CAM. The experimental results show that DAM-EfficientNet-B0 accomplishes accuracy of 99.21% on EyePACS and accuracy of 99.47% on Messidor datasets, outperforming the existing models: MS-DNN and DNN-BOA. In the future work, improved DL methods with optimization algorithms and transformer-based approaches can be implemented to enhance the detection classification of DR.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper background work, conceptualization, methodology, dataset collection, have been done by 1st and 5th author. Implementation, result analysis and comparison have been done by 3rd and 4th author. Preparing and editing draft, visualization, supervision, review of work and project administration have been done by 2nd Author.

Notation

Notation	Description
g and g_{min}	Measured and minimum pixel values
β	Threshold parameter
$p(F)$	Distribution function
$y_{i,n} - x_i^{(t)}$	Finite Gaussian mixture model
$\alpha_{i,j}$	Gross error mode probability
C	Relevant feature in spatial
H	Spatial Dimension height
W	Spatial dimension width
F	Input feature map
F1, F2, and F3	New feature map
$H \times W$	Scale of feature map
W_1 and W_2	Weight of fully connected
σ	Sigmoid activation
U_{avg} and U_{max}	Average and Maximum Pooling Feature Map
$w, d,$ and r	Network's width, depth, and resolutions
α, β and γ	Constant terms of hyper parameter range

References

[1] M.S.B. Phridviraj, R. Bhukya, S. Madugula, A. Manjula, S. Vodithala, and M.S. Waseem, "A bi-directional Long Short-Term Memory-based Diabetic Retinopathy detection model using retinal fundus images", *Healthcare Analytics*, Vol. 3, p. 100174, 2023.

[2] V.D. Vinayaki, and R. Kalaiselvi, "Multithreshold image segmentation technique using remora optimization algorithm for diabetic retinopathy detection from fundus images", *Neural Processing Letters*, Vol. 54, pp. 2363-2384, 2022.

[3] D. Das, S.K. Biswas, and S. Bandyopadhyay, "Detection of diabetic retinopathy using convolutional neural networks for feature extraction and classification (DRFEC)", *Multimedia Tools and Applications*, Vol. 82, No. 19, pp. 29943-30001, 2023.

[4] G. Ali, A. Dastgir, M.W. Iqbal, M. Anwar, and M. Faheem, "A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images", *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 11, pp. 341-350, 2023.

[5] M.M. Butt, D.N.F.A. Iskandar, S.E. Abdelhamid, G. Latif, and R. Alghazo, “Diabetic retinopathy detection from fundus images of the eye using hybrid deep learning features”, *Diagnostics*, Vol. 12, No. 7, p. 1607, 2022.

[6] A. Jabbar, H.B. Liaqat, A. Akram, M.U. Sana, I.D. Azpíroz, I.D.L.T. Diez, and I. Ashraf, “A Lesion-Based Diabetic Retinopathy Detection Through Hybrid Deep Learning Model”, *IEEE Access*, Vol. 12, pp. 40019-40036, 2024.

[7] F.M.J.M. Shamrat, R. Shakil, Sharmin, N. Hoque ovy, B. Akter, M.Z. Ahmed, K. Ahmed, F.M. Bui, and M.A. Moni, “An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection”, *Healthcare Analytics*, Vol. 5, p. 100303, 2024.

[8] M.H. Mahmoud, S. Alamery, H. Fouad, A. Altinawi, and A.E. Youssef, “An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm”, *Personal and Ubiquitous Computing*, Vol. 27, No. 3, pp. 751-765, 2023.

[9] B. Venkaiahppalaswamy, P.V.G.D.P. Reddy, and S. Batha, “Hybrid deep learning approaches for the detection of diabetic retinopathy using optimized wavelet based model”, *Biomedical Signal Processing and Control*, Vol. 79, No. 2, p. 104146, 2023.

[10] F. Mercaldo, M. Di Giammarco, A. Apicella, G. Di Iadarola, M. Cesarelli, F. Martinelli, and A. Santone, “Diabetic retinopathy detection and diagnosis by means of robust and explainable convolutional neural networks”, *Neural Computing and Applications*, Vol. 35, No. 23, pp. 17429-17441, 2023.

[11] B. Menaouer, Z. Dermane, N. El Houda Kebir, and N. Matta, “Diabetic retinopathy classification using hybrid deep learning approach”, *SN Computer Science*, Vol. 3, p. 357, 2022.

[12] S.J. Minija, M.A. Rejula, and B.S. Ross, “Automated detection of diabetic retinopathy using optimized convolutional neural network”, *Multimedia Tools and Applications*, Vol. 83, No. 7, pp.21065-21080, 2024.

[13] E. Özbay, “An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm”, *Artificial Intelligence Review*, Vol. 56, pp. 3291-3318, 2023.

[14] Y. Chen, S. Xu, J. Long, and Y. Xie, “DR-Net: Diabetic retinopathy detection with fusion multi-lesion segmentation and classification”, *Multimedia Tools and Applications*, Vol. 82, No. 17, pp. 26919-26935, 2023.

[15] A. Erciyas, N. Barışçı, H.M. Ünver, and H. Polat, “Improving detection and classification of diabetic retinopathy using CUDA and Mask RCNN”, *Signal, Image and Video Processing*, Vol. 17, No. 4, pp. 1265-1273, 2023.

[16] H. Mustafa, S.F. Ali, M. Bilal, and M.S. Hanif, “Multi-stream deep neural network for diabetic retinopathy severity classification under a boosting framework”, *IEEE Access*, Vol. 10, pp. 113172-113183, 2022.

[17] D.A. Da Rocha, F.M.F. Ferreira, and Z.M.A. Peixoto, “Diabetic retinopathy classification using VGG16 neural network”, *Research on Biomedical Engineering*, Vol. 38, No. 2, pp. 761-772, 2022.

[18] S.R. Syed, and S.M.A. Durai, “A diagnosis model for detection and classification of diabetic retinopathy using deep learning”, *Network Modeling Analysis in Health Informatics and Bioinformatics*, Vol. 12, p. 37, 2023.

[19] G. Kalyani, B. Janakiramaiah, A. Karuna, and L.V.N. Prasad, “Diabetic retinopathy detection and classification using capsule networks”, *Complex & Intelligent Systems*, Vol. 9, No. 3, pp. 2651-2664, 2023.

[20] V. Rachapudi, K.S. Rao, T.S.M. Rao, P. Dileep, and T.L.D. Roy, “Diabetic retinopathy detection by optimized deep learning model”, *Multimedia Tools and Applications*, Vol. 82, No. 18, pp. 27949-27971, 2023.

[21] P. Saranya, R. Pranati, and S.S. Patro, “Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models”, *Multimedia Tools and Applications*, Vol.82, No. 25, pp. 39327-39347, 2023.

[22] EyePACS Dataset: https://paperswithcode.com/dataset/kaggle-eyepacs (Accessed on June 2024).

[23] Messidor Dataset: https://paperswithcode.com/dataset/messidor-1 (Accesses on June 2024).

[24] G.U. Nneji, J. Cai, J. Deng, H.N. Monday, M.A. Hossin, and S. Nahar, “Identification of Diabetic Retinopathy Using Weighted Fusion Deep Learning Based on Dual-Channel Fundus Scans”, *Diagnostics*, Vol. 12, p. 540, 2022.

[25] T. Prabhakar, T.V.M. Rao, B. Maram, and D. Chigurukota, “Exponential gannet firefly

optimization algorithm enabled deep learning for diabetic retinopathy detection", *Biomedical Signal Processing and Control*, Vol. 87A, p. 105376, 2024.

[26] M. Zubair, M. Umair, R.A. Naqvi, D. Hussain, M. Owais, and N. Werghi, "A comprehensive computer-aided system for an early-stage diagnosis and classification of diabetic macular edema", *Journal of King Saud University-Computer and Information Sciences*, Vol. 35, No. 8, p. 101719, 2023.

[27] S.H. Abbood, H.N.A. Hamed, M.S.M. Rahim, A. Rehman, T. Saba, and S.A. Bahaj, "Hybrid retinal image enhancement algorithm for diabetic retinopathy diagnostic using deep learning model", *IEEE Access*, Vol. 10, pp. 73079-73086, 2022.