[1] Bhavani HS
[2] Dr. Karunakara K.

Otsu Multi-level Adaptive Thresholding with Deep Convolution Neural Network for Diabetic Retinopathy Segmentation and Classification

Journal of Electrical Systems

Abstract: - Diabetic Retinopathy (DR) is eye disease caused through the retinal damage induced through long-term diabetes mellitus illness. In the initial stage, DR has minor vision difficulties or no symptoms but it outputs in vision loss if it is not treated early. However, overfitting occurs when a model learns noise or specific features from a small dataset which leading to poor generalization. To solve this issue, this research proposed an Otsu Multi-level Adaptive Thresholding with Deep Convolution Neural Network (OMAT+DCNN) approach for DR segmentation and classification. The Contrast Limited Adaptive Histogram Equalization (CLAHE) is used for preprocessing input image which enhances the contrast level and image visibility without amplifying noise. The OMAT is used for segmenting the affected regions and ResNet50 is used for extracting Deep Learning (DL) features. Then, relevant features are selected through Multi-Verse Optimizer (MVO) and DCNN is used for classification. The OMAT improves segmentation accuracy by considering multiple thresholds which adopts to local intensity variations within the image, even when the dataset is limited. DCNN capturing intricate patterns and spatial hierarchies within images and its deep architecture allows to learn complex representations which leads to enhance accuracy. The parameters like dice coefficient, Intersection over Union (IoU), Mean IoU (MIoU), precision, accuracy, f1-score and sensitivity are used for assessing performance. The OMAT+DCNN reaches 0.993 of accuracy when compared to existing techniques such as Convolutional Transformer Network (CTNet) and Residual Attention Network (RAN).

Keywords: Contrast Limited Adaptive Histogram Equalization, Deep Convolution Neural Network, Diabetic Retinopathy, Multi-Verse Optimizer, Otsu Multi-level Adaptive Thresholding

I. INTRODUCTION

The Diabetic Retinopathy (DR) is a blinding eye disease related with long-term diabetes which has major cause of blindness in worldwide [1]. It is one of the significant difficulties of diabetes and analysis denotes that 1/3 of diabetic patients will grow DR, that is also major concern for enhancing the blindness count [2]. The blood vessels damage the light-sensitive tissue called retina that is placed in the back of eye which leads disorder [3]. The dilated eye examination is known as fluorescein angiography which is conducted for DR diagnosis in patient and it demands the occurrence of clinical experts and ophthalmologists [4]. The best alternatives are required due to the present approach relies on resources which is time-consuming. The DR presents with no symptoms which leads vision problem initially and significant visual damage [5]. The early consultation of patient with diabetes, experts for DR evaluation and assessment has required which minimize the DR growth and susceptibility to severe blindness [6]. Whether the DR remains undiagnosed which growths in a procedure of different stages with lesion formation like Exudates (EXs), Hemorrhages (HEs), Microaneurysms (Mas), abnormal Foveal Avascular Zone (FAZ), Intra Retinal Microvascular Abnormalities (IRMAs), Optic Disc (OD) abnormal structure and Neovascularization and so on [7, 8]. The domain experts analyze the human retina through high-resolution digitized fundus camera for acquiring fundus images [9].

In the DR diagnosis process, clinical experts perform detailed manual analysis for fundus image to recognize retinal features which is not only need substantive human intervention and huge period however its prone human error or bias which is harmful for early detection [10, 11]. The automated technique applications for resolving issues of science and technology has enhanced in recent past years [12]. The appearance of Deep Learning (DL) and Machine Learning (ML) in image analysis, segmentation, classification and so on [13]. To address DR diagnosis issues providing to higher population at speed and enhance the performance and automated screening methods are developed [14]. This minimizes the vision abnormality risk between asymptomatic person affected through DR and warns patient to pursue instant treatment to cure impairment issue or vision loss [15, 16]. The DR denotes retinal abnormalities with venous bleeding, shrinking or swelling of blood vessels, existing blood vessels dilation, abnormal branches or development of new vessels [17]. However, overfitting occurs when a model learns noise or

[1] Research Scholar, SSAHE, Assistant Professor, Department of MCA SSIT, Tumkur, Karnataka, India. Email: bhavanihs@ssit.edu.in

[2] Research Supervisor, SSAHE,Tumkur , Professor, Department of ISE SSIT, Tumkur, Karnataka, India.

specific features from a small dataset which leading to poor generalization. This research proposed an OMAT+DCNN in which the OMTC enhances the segmentation accuracy through considering multiple thresholds and DCNN captures complex patterns and spatial hierarchies in the image which leads high accuracy. The main contribution of proposed research are as follows:

- The OMAT improves segmentation accuracy by considering multiple thresholds which adopts to local intensity variations within the image, even when the dataset is limited. By segmenting regions based on intensity levels, it provides more meaningful information for subsequent analysis.
- The OMAT avoids overfitting by focusing on intensity-based segmentation which doesn't rely on complex model parameters or feature extraction. By emphasizing local intensity variations, it captures relevant structures without overfitting to noise.
- DCNN capturing intricate patterns and spatial hierarchies within images and its deep architecture allows them to learn complex representations, leading to improved accuracy. Deep layers capture low-level features (edges, textures), while higher layers combine these features to recognize more complex structures (lesions, blood vessels). This hierarchical approach enables better discrimination between different disease stages.

The structure of manuscript is as follows: The existing work is described in Section 2, details of proposed method is provided in Section 3, result and discussion are given in Section 4, conclusion and future scope are presented in Section 5.

II. LITERATURE REVIEW

Numerous research has been developed in the automated DR segmentation and classification area with various DL techniques. The continuous activities have been considered to enhance the automated tool effectiveness to serve the population through world and aid health practitioners.

Douglas Abreu da Rocha et al. [18] suggested a VGG16 neural network-based DR classification. The suggested approach involves retinal image preprocessing, consisting size adequacy, data cleaning, augmentation and class imbalance at training stage. The hyperparameter tuning and image classification through VGG16 network. The suggested approach easy to handle large-dimensional data which makes it appropriate for resolving difficult problems. However, it prone to overfitting issues particularly dealing with noisy data.

Ruchika Bala et al. [19] implemented a CTNet for DR classification. It includes the integration of CNN and transformer to capture global and local features. Initially, convolution model was designed through residual connection for local lesion feature extraction. Then, transformer model patch features to small patch of sequence which defines global long-range focused one lesion area through self-attention. Atlast, pooling was accomplished on sequence patches through memory-inefficient class tokens for produce classification. It learns meaningful and discriminative features from input which leads the model more effective however, it was sensitive to noisy data and tried to recognize optimal hyperplane.

Moye Yu and Yi Wang [20] introduced a RAN based DR detection and classification. Initially, the fundus images are preprocessed through normalization, enhancement which was provided to CNN for obtain classification result. The RAN was developed according to residual attention which integrated dilated convolution to attain optimal result. Atlast, the focal loss was included to resolve class imbalance issue. It capable to handle long-term dependencies and sequential information from input data however, it struggles to learn certain features and incapability to manage temporal dependencies which makes the model complex to train.

Paresh Chandra Sau and Atul Bansal [21] presented a DR grading through Modified Deep Neural Network (MDNN) with blood vessels and retinal abnormality segmentation. Initially, median filter was used for preprocessing and before starting segmentation, optic disc was eliminated and adaptive action contour was applied for segmentation. Then, MCNN was optimized through Fitness-based Newly Updated Grasshopper Optimization Algorithm (FNU-GOA). The FNU-GOA-MDNN was effectively segmented and classified the DR which achieved better classification accuracy. However, it was not adoptable to handle high-dimensional input due to local optima issue.

P Saranya et al. [22] introduced a DR detection and classification through CNN which classify the input with its respective classes. The preprocessing includes to remove noise, enhance local contrast level which is segmented through U-Net. The advanced convolution layer in U-Net was applied to support pixel-level labelling that was required in segmentation. The developed model concentrated on significant regions according to disease diagnosis. However, the irrelevant features are not selected through feature selection which resultant in less classification accuracy.

From the above analysis, various DL approaches are developed in past years which has drawbacks. It is prone to overfitting issues particularly dealing with noisy data, it was sensitive to noisy data and tried to recognize optimal hyperplane. It struggles to learn certain features and incapability to manage temporal dependencies which makes the model complex to train. It was not adoptable to handle high-dimensional input due to local optima issue. The irrelevant features are not selected through feature selection which resultant in less classification accuracy. Based on this analysis, this research proposed an OMAT-DCNN for segmentation and classification of DR.

III. PROPOSED METHOD

This research proposed an Otsu Multi-level Adaptive Thresholding with Deep Convolution Neural Network (OMAT+DCNN) approach for DR segmentation and classification. The IDRID dataset is preprocessed by CLAHE which enhances the contrast level and image visibility. The OMAT is used for segmenting the affected regions and ResNet50 is used for extracting DL features. Then, relevant features are selected through MVO and DCNN is used for classification. Figure 1 denotes the schematic presentation of proposed OMAT-DCNN approach.

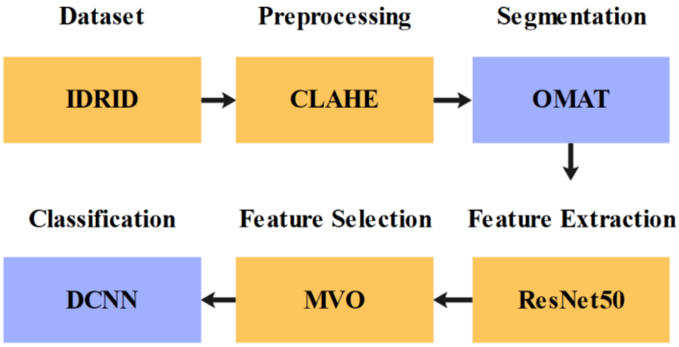

Figure 1. Schematic presentation of proposed OMAT-DCNN approach

3.1 Dataset

The IDRID (Indian Diabetic Retinopathy Image Dataset) is a public dataset which contains 516 images [23] labeled for DR classification. In that 516 images, 413 images (80%) are provided for training and 103 images (20%) are provided for testing. These 516 retinal fundus images have high resolution of 4288×2848 pixels. Figure 2 denote the dataset sample image.

Figure 2. Dataset sample image

3.2 Preprocessing

The dataset is preprocessed through CLAHE which enhances the contrast level and image visibility without amplifying noise. The CLAHE is used for preprocessing the input image, which subsequently applied to the denoised images of retinal fundus to enhance the contrast. It is a sophisticated variety of Adaptive Histogram Equalization (AHE). After CLAHE, the output is shown in eq. (1),

$$CLAHE = \{f(Z_m(x,y)) | \forall Z_m(x,y) \in Z_m\} \tag{1}$$

Here, $f(Z_m)$ is a transformation-based density function, (x,y) is a coordinates of pixel image. It enhances the contrast level from input image and assist to evade saturation and clipping problem which enhance the image visibility without amplifying noise. Figure 3 denotes the preprocessed image.

Figure 3. Preprocessed image

3.3 Segmentation

The preprocessed image is given into segmentation process which segments the affected regions through OMAT approach. Otsu's method uses the grayscale histogram of an image to find an optimal threshold value that separates the foreground and background regions with maximum inter-class variance and it automatically determines the threshold by maximizing the separation between pixel intensities, making it robust and adaptive. In multi-level thresholding, image with gray level is separated into various regions in which the images are separated into regions with brightness which is related to various objects. In this approach, complex action selects suitable threshold value automatically and numerous algorithms unable to attain above scores repeatedly. One or more regions defines objects from images due to this, it used for color images. The R, G, and B channels independently processed and formulated through eq. (2),

$$\begin{aligned} T_0 &= \{I(x.y) \in X | 0 \leq I(x.y) \leq t_1 - 1\} \\ T_1 &= \{I(x.y) \in X | t_1 \leq I(x.y) \leq t_2 - 1\} \\ T_i &= \{I(x.y) \in X | t_i \leq I(x.y) \leq t_{i+1} - 1\} \\ T_m &= \{I(x.y) \in X | t_m \leq I(x.y) \leq N - 1\} \end{aligned} \tag{2}$$

Where, x is image which is processed, $I(x.y)$ is an intensity score of pixel responder that is denoted through coordinate scores indicated as x and y, $t_i = 1,2, \dots, m$. The Otsu uses histogram in an image and determine ideal global threshold. Otsu thresholding is used in this research on retinal fundus images to separate background and Region of Interest (ROI) of an image. It is a nonlinear transformation of a grayscale image to a binary image. The algorithm's input is a grayscale image, and its output will be a binary image that has only same pixel as input's image intensity. If pixel intensity of input is above threshold, output pixel value will be white. Otherwise, if intensity of input pixel is less than threshold or equal, output pixel value is zero or black. The Otsu technique is used to enhance the variance among image class I as signified in eq. (3), (4),

$$F_{otsu} = \sum_{i=0}^{K} \theta_i \times (\mu_i - \mu_1)^2 \cdot \theta_i = \sum_{j=t_i}^{t_{i+1}-1} P_j \tag{3}$$

$$\mu_i = \sum_{j=t_i}^{t_{i+1}-1} \frac{iP_j}{\theta_j} \, .P_i = \frac{Fr_i}{N_P} \cdot t_0 = 0 \, \cdot \, t_{K+1} = L \tag{4}$$

Where, P_i and Fr_i are possibility and frequency of gray level in ith image, μ_1 is an average image intensity i, N_P is a total number of image pixel. The multilevel thresholding is a prominent thresholding technique, which defines threshold from divided groups according to entropy. The same rule is used for multi-level threshold which is numerically signified in eq. (5),

$$H_0 = -\sum_{i=0}^{t_1-1}\left(\frac{P_i}{\omega_0}\right)\log_2\left(\frac{P_i}{\omega_0}\right);\ H_1 = -\sum_{i=t_1}^{t_2-1}\left(\frac{P_i}{\omega_1}\right)\log_2\left(\frac{P_i}{\omega_1}\right);$$
$$H_j = -\sum_{i=t_j}^{t_{j+1}-1}\left(\frac{P_i}{\omega_{0j}}\right)\log_2\left(\frac{P_i}{\omega_j}\right);\ H_m = -\sum_{i=t_m}^{N-1}\left(\frac{P_i}{\omega_m}\right)\log_2\left(\frac{P_i}{\omega_m}\right); \tag{5}$$

Where, $\omega_0 = \sum_{i=0}^{t_1-1} P_i$; $\omega_1 = \sum_{i=0}^{t_2-1} P_i$; $\omega_j = \sum_{i=t_j}^{t_{j+1}-1} P_i$; $\omega_m = \sum_{i=t_m}^{N-1} P_i$; $H_0, H_1 \dots H_m$ are entropy scores of $m + 1$ of various regions, N is a total number of intensity level in grayscale image, P_i is a pixel intensity probability. It segments the color images with R, G and B channels processed individually. Figure 4 denotes the segmented image.

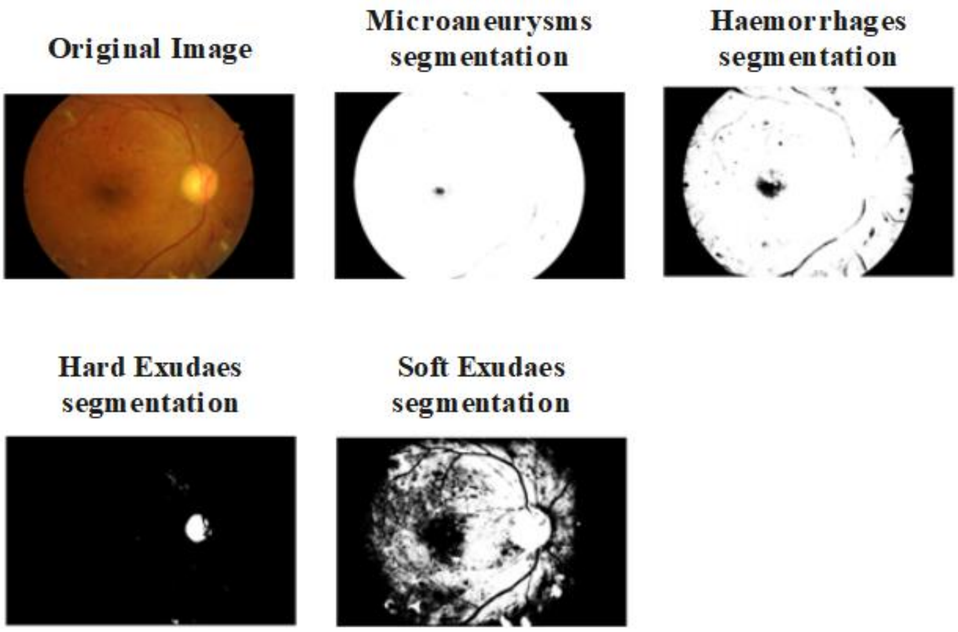

Figure 4. Segmented images

3.4 Feature Extraction

The segmented images are given to feature extraction for extracting deep learning features. The ResNet50 is used to extract features from segmented images to aid accurate detection of DR. Here, the ResNet50 extracts 2048 features. The residual block increases network potency and performance. It is capable of generating the best output in classifications. The residual block implements a residual below the inclusion of residual and replication. The extracted features are selected by feature selection process.

3.5 Feature Selection

The Multi-Verse Optimizer (MVO) is used to selected 1426 relevant features from the extracted features. The MVO is a metaheuristic algorithm which is inspired through famous theory named as Multi-Verse theory. This theory was developed according to three cosmologic concepts such as black holes, white holes and wormholes. The white holes have strong repulsion and release every objects; black holes have high gravitational forces extremely and it fascinates every objects; wormholes join various universes and transported object orbits. The MVO working process is according to the principle which matter in the universe transfer from white and black holes in wormholes. Based on the integrated actions of white, black and wormholes, the whole multiverse population reach convergence state eventually. MVO classifies the searching process in two stages: exploration and exploitation. The exploration is done by exchange of white/black holes and exploitation procedure is done by wormholes. The MVO iterates primary universe regularly by white/black holes in that universe denotes the feasible solution to the issues and the universe object denotes components of solution, the universe response rate denotes fitness value of each solution. It is numerically signified as follows:

Multiverse Initialization: Initially, a set of random universes U is created by using eq. (6),

$$U = \begin{bmatrix} x_1^1 & x_1^2 & \cdots & x_1^d \\ x_2^1 & x_2^2 & \cdots & x_2^d \\ \vdots & \vdots & \vdots & \vdots \\ x_n^1 & x_n^2 & \cdots & x_n^d \end{bmatrix} \tag{6}$$

Where, d is a variable, n is a number of universes (candidate solution).

Black and white hole mechanisms: Because of various expansion rate of every individual universe, objects in this universe are transported into black/white holes. This procedure follows roulette mechanism as signified in eq. (7),

$$x_i^j = \begin{cases} x_k^j & r_1 < NI(U_i) \\ x_i^j & r_1 \geq NI(U_i) \end{cases} \tag{7}$$

Where, x_i^j is a ith variable in jth universe, $NI(U_i)$ is a normalized expansion rate of universe i, r_1 is a random number among $[0, 1]$, x_k^j is a ith variable in jth universe chosen based on spiral mechanism.

Wormhole mechanism: Without the expansion rate size, to achieve local changes and enhance its expansion rate, the individual universe is inspired based on internal objects to transfer present optimal universe as signified in eq. (8), (9),

$$WEP = WEP_{min} + l \times \left(\frac{WEP_{min} - WEP_{max}}{L}\right) \tag{8}$$

$$TDR = 1 - \frac{l^{1/p}}{L^{1/p}} \tag{9}$$

Where, WEP is a possibility of wormholes existence in multiverse space, WEP_{min} and WEP_{max} are minimum and maximum possibilities, l and L are present and maximum iterations, TDR is a numerical value of travel distance, p is a development accuracy in each iteration. The high score of p, quicken the local development but minimize the search scope which is specified in eq. (10),

$$x_i^j = \begin{cases} \begin{cases} X_j + TDR \times \left(\left(ub_j - lb_j\right) \times r_4 + lb_j\right), r_3 < 0.5, \\ X_j + TDR \times \left(\left(ub_j - lb_j\right) \times r_4 + lb_j\right), r_3 \geq 0.5, \end{cases} & r_2 < WEP \\ x_i^j & r_2 \geq WEP \end{cases} \tag{10}$$

Where, x_i^j is a jth object of present optimal universe, ub_j and lb_j are lower and upper bounds of x_i^j, r_2, r_3 and r_4 are random numbers among $[0, 1]$.

3.6 Classification

The selected features are used for the classification of DR through DCNN model which classifies into four classes Microaneurysms, haemorrhages, hard exudaes and soft exudaes. DCNNs excel in capturing intricate patterns and spatial hierarchies within images. Their deep architecture allows them to learn complex representations, leading to improved accuracy. DCNN learn hierarchical representations of visual patterns. Deep layers capture low-level features (edges, textures), while higher layers combine these features to recognize more complex structures (lesions, blood vessels). This hierarchical approach enables better discrimination between different disease stages. The convolution, maxpooling and dense layer are CNN components. The Convolution layers collect features from input by filters and maxpooling is utilized to low layer size which enhances the computational effectiveness. The dense layer helps to connect layers which is denotes as Fully Connected (FC) layer.

Input layer: By this layer, model communicate the data which was received at input to hidden layer and forward to output layer.

Convolution layer: It is a basic component of convolution network that has few parameters which are made up of filters, by these filters the layer pulls and learn features. The input is compared by segments to define variations and these segments are denotes as features. This layer extracts features and utilizes matrix to build matrices and dot product. Additionally, the entire process is output denoted as convolution layer. For input (X) and kernel K, 2D-convolution operator is signified in eq. (11),

$$(X \times K)(i,j) = \sum_{m}\sum_{n} K(m,n)X(i-m,j-n) \tag{11}$$

Where, $\times$ is a mathematical presentation of convolution operation.

Pooling layer: It is next layer after convolution layer which is generally applied to formed feature map to minimize the feature map and network parameters through applying respective mathematical computation. The max pooling selects one maximum score through matrix size denoted in every map which outputs in reduced neurons. The global average pooling used before FC which reduce data to one-dimension. The dropout layer is middle layer which prevents network from overfitting.

Batch Normalization Layer: It is used to address the issue of internal covariant shift that occurs when layer parameters change prior layer parameters at training. The normalization through batch among layers allows to attain high learning rate which outputs as quicken training process.

Dropout Layer: It is regularization technique which minimizes the complexity and assist to resolve overfitting issue. At the training of neural network, it eliminates the units randomly through setting activation layers to 0.

Dense Layer: It is responsible for defining connection among every characteristic which provided to it without the usage of input parameters rather than convolution layers.

Output layer: It is responsible for obtaining final class due to its sigmoid function it works better when classifying two classes. The softmax is applied for multi-classification which ensures total possibilities of result. At result, it is considered softmax activation function is selected for proposed model which is signified in eq. (12),

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{i=1}^{K} e^{z_k}} \; fori = 0,1,\ldots,k \tag{12}$$

This function receives K dimensional vector z and returns K dimensional vector comprising values in the range of [0,1].

IV. EXPERIMENTAL RESULT

In this manuscript, OMAT+DCNN is simulated in python environment with system requirement of 8GB RAM, intel core i5 processor and windows 10 OS. The parameters like dice coefficient, IoU, MIoU, precision, accuracy, f1-score and sensitivity are used for assessing performance which is numerically signified in eq. (13-19),

$$Dice\ coefficient = \frac{2 \times TP}{2 \times TP + FP + FN} \times 100 \tag{13}$$

$$IoU = \frac{TP}{TP + FP + FN} \times 100 \tag{14}$$

$$MIoU = \frac{1}{N}\sum_{i=1}^{N} IoU \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$F - Measure = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity} \tag{18}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{19}$$

Where, TP, TN, FP and FN signifies True Positive, True Negative, False Positive and False Negative, N is a number of classes.

4.1 Quantitative and Qualitative Analysis

This section presents the performance of OMAT+DCNN in IDRID dataset in terms of segmentation and classification. The dice coefficient, IoU, MIoU, precision, accuracy, f1-score and sensitivity are considered as performance metrics. Table1 shows the segmentation result, table 2 shows the classification result with actual and optimized features. Figure 5 and 6 denotes the epoch vs accuracy and epoch vs loss graph.

Table 1. Result of Segmentation in IDRID dataset

Method	**Dice coefficient**	**IoU**	**MIoU**
Otsu thresholding	0.975	0.925	0.936
Multi-level	0.964	0.954	0.961
Adaptive multi-level	0.968	0.970	0.965
OMAT	**0.988**	**0.985**	**0.972**

Table 1 exemplifies the OMAT performance in IDRID dataset in terms of dice coefficient, IoU, and MIoU metrics. The Otsu thresholding, multi-level thresholding, adaptive multi-level thresholding approaches are considered to compare with proposed OMAT approach. The OMAT realizes 0.988 of dice coefficient, 0.985 of IoU, and 0.972 of MIoU which is better compared to existing segmentation approaches.

Table 2. Result of Classification in IDRID dataset

Actual Features				
Method	**Precision**	**Accuracy**	**F1-score**	**Sensitivity**
RNN	0.919	0.945	0.957	0.958
DNN	0.967	0.959	0.973	0.963
CNN	0.972	0.966	0.979	0.971
DCNN	**0.978**	**0.973**	**0.981**	**0.977**
Optimized Features				
Method	**Precision**	**Accuracy**	**F1-score**	**Sensitivity**
RNN	0.965	0.956	0.950	0.954
DNN	0.969	0.960	0.969	0.976
CNN	0.986	0.979	0.986	0.986
DCNN	**0.999**	**0.993**	**0.997**	**0.991**

Table 2 exemplifies the DCNN performance in IDRID dataset in terms of precision, accuracy, f1-score and sensitivity metrics. The Recurrent Neural Network (RNN), Deep Neural Network (DNN), and CNN are considered to compare with proposed DCNN approach. The DCNN realizes 0.978 of precision, 0.973 of accuracy, 0.981 of f1-score, 0.977 of sensitivity for actual features; 0.999 of precision, 0.993 of accuracy, 0.997 of f1-score, 0.991 of sensitivity for optimized features which is better compared to existing classifiers. Figure 5 and 6 signifies the epoch vs accuracy and epoch vs loss graph for IDRID dataset.

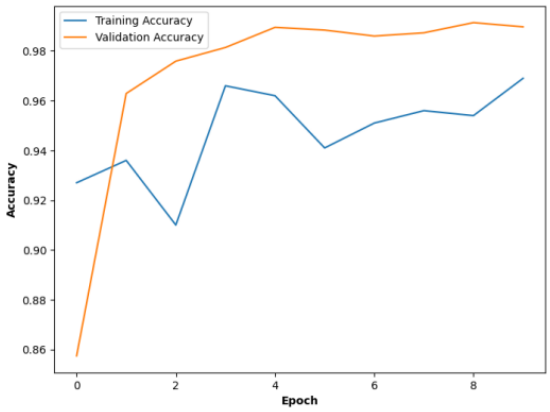

Figure 5. Epoch vs Accuracy

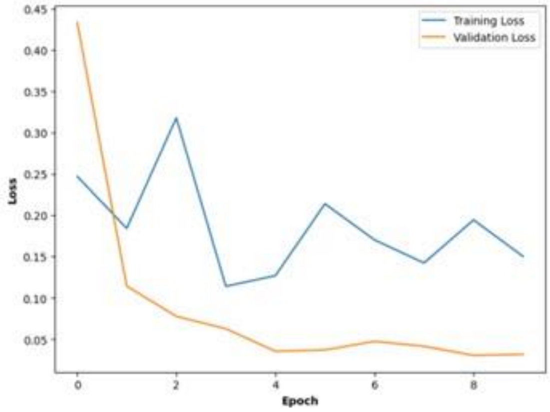

Figure 6. Epoch vs Loss

4.2 Comparative Analysis

The proposed OMAT-DCNN performance is compared with existing approaches such as VGG16 [18], CTNet [19], RAN [20] and FNU-GOA-MDNN [21]. The precision, accuracy, f1-score and sensitivity are used for assessing performance of OMAT-DCNN approach. The OMAT+DCNN realizes 0.999 of precision, 0.993 of accuracy, 0.997 of f1-score, 0.991 of sensitivity for IDRID dataset. Table 3 shows the comparative analysis in IDRID dataset.

Table 3. Comparative analysis in IDRID dataset

Method	**Precision**	**Accuracy**	**F1-score**	**Sensitivity**
VGG16 [18]	0.850	0.898	0.877	0.868
CTNet [19]	0.990	0.989	0.991	0.989
RAN [20]	NA	0.715	NA	0.752
FNU-GOA-MDNN [21]	0.984	0.952	0.975	0.975
OMAT+DCNN	**0.999**	**0.993**	**0.997**	**0.991**

4.3 Discussion

The existing DR techniques such as VGG16 [18] prone to overfitting issues particularly dealing with noisy data. CTNet [19] was sensitive to noisy data and tried to recognize optimal hyperplane. RAN [20] struggles to learn certain features and incapability to manage temporal dependencies which makes the model complex to train. FNU-GOA-MDNN [21] not adoptable to handle high-dimensional input due to local optima issue. To solve these problem, this research proposed an OMAT-DCNN overcomes the overfitting issues. The OMAT improves segmentation accuracy through taking multiple thresholds that adopts to local intensity variations in image. The DCNN captures complex patterns and spatial hierarchies within images and its deep architecture allows to learn complex representations that's leads to enahnce accuracy.

V. CONCLUSION

This research proposed an Otsu Multi-level Adaptive Thresholding with Deep Convolution Neural Network (OMAT+DCNN) approach for DR segmentation and classification. The CLAHE is used for preprocessing input image which enhances the contrast level and image visibility without amplifying noise. The OMAT is used for segmenting the affected regions and ResNet50 is used for extracting DL features. Then, relevant features are selected through MVO and DCNN is used for classification. The OMAT enhances segmentation accuracy through considering multiple thresholds that adopts to local intensity variations within the image, even it has limited dataset. By segmenting regions based on intensity levels, it provides more meaningful information for subsequent analysis. The DCNN captures intricate patterns and spatial hierarchies within images and its deep architecture enables to learn complex representations which leads to improved accuracy. The OMAT+DCNN reaches 0.999 of precision, 0.993 of accuracy, 0.997 of f1-score, 0.991 of sensitivity for IDRID dataset. In future various DL methods can be utilized for enhancing the classification performance.

REFERENCES

[1] Das, D., Biswas, S.K. and Bandyopadhyay, S., 2023. Detection of diabetic retinopathy using convolutional neural networks for feature extraction and classification (DRFEC). Multimedia Tools and Applications, 82(19), pp.29943-30001.

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

[3] Srinivasan, V. and Rajagopal, V., 2022. Multi-Scale Attention-Based Mechanism in Gradient Boosting Convolutional Neural Network for Diabetic Retinopathy Grade Classification. International Journal of Intelligent Engineering & Systems, 15(4).

[4] Sambyal, N., Saini, P., Syal, R. and Gupta, V., 2023. Modified residual networks for severity stage classification of diabetic retinopathy. Evolving Systems, 14(1), pp.17-35.

[5] Zhang, G., Sun, B., Zhang, Z., Pan, J., Yang, W. and Liu, Y., 2022. Multi-model domain adaptation for diabetic retinopathy classification. Frontiers in Physiology, 13, p.918929.

[6] Raiaan, M.A.K., Fatema, K., Khan, I.U., Azam, S., Rashid, M.R.U., Mukta, M.S.H., Jonkman, M. and De Boer, F., 2023. A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images. IEEE Access, 11, pp.42361-42388.

[7] Fang, L. and Qiao, H., 2023. A novel DAG network based on multi-feature fusion of fundus images for multi-classification of diabetic retinopathy. Multimedia Tools and Applications, 82(30), pp.47669-47693.

[8] Mehboob, A., Akram, M.U., Alghamdi, N.S. and Abdul Salam, A., 2022. A Deep Learning Based Approach for Grading of Diabetic Retinopathy Using Large Fundus Image Dataset. Diagnostics, 12(12), p.3084.

[9] Chaudhary, P.K. and Pachori, R.B., 2022. Automatic diagnosis of different grades of diabetic retinopathy and diabetic macular edema using 2-D-FBSE-FAWT. IEEE Transactions on Instrumentation and Measurement, 71, pp.1-9.

[10] Bilal, A., Imran, A., Baig, T.I., Liu, X., Long, H., Alzahrani, A. and Shafiq, M., 2024. DeepSVDNet: A Deep Learning-Based Approach for Detecting and Classifying Vision-Threatening Diabetic Retinopathy in Retinal Fundus Images. Computer Systems Science & Engineering, 48(2).

[11] Albadr, M.A.A., Ayob, M., Tiun, S., Al-Dhief, F.T. and Hasan, M.K., 2022. Gray wolf optimization-extreme learning machine approach for diabetic retinopathy detection. Frontiers in Public Health, 10, p.925901.

[12] Atteia, G., Abdel Samee, N., El-Kenawy, E.S.M. and Ibrahim, A., 2022. CNN-hyperparameter optimization for diabetic maculopathy diagnosis in optical coherence tomography and fundus retinography. Mathematics, 10(18), p.3274.

[13] Pavani, P.G., Biswal, B. and Gandhi, T.K., 2023. Simultaneous multiclass retinal lesion segmentation using fully automated RILBP-YNet in diabetic retinopathy. Biomedical Signal Processing and Control, 86, p.105205.

[14] Vij, R. and Arora, S., 2023. A novel deep transfer learning based computerized diagnostic Systems for Multi-class imbalanced diabetic retinopathy severity classification. Multimedia Tools and Applications, 82(22), pp.34847-34884.

[15] Shaukat, N., Amin, J., Sharif, M., Azam, F., Kadry, S. and Krishnamoorthy, S., 2022. Three-dimensional semantic segmentation of diabetic retinopathy lesions and grading using transfer learning. Journal of Personalized Medicine, 12(9), p.1454.

[16] Skouta, A., Elmoufidi, A., Jai-Andaloussi, S. and Ouchetto, O., 2022. Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network. Journal of Big Data, 9(1), p.78.

[17] Zia, F., Irum, I., Qadri, N.N., Nam, Y., Khurshid, K., Ali, M., Ashraf, I. and Khan, M.A., 2022. A multilevel deep feature selection framework for diabetic retinopathy image classification. Comput. Mater. Contin, 70, pp.2261-2276.

[18] Da Rocha, D.A., Ferreira, F.M.F. and Peixoto, Z.M.A., 2022. Diabetic retinopathy classification using VGG16 neural network. Research on Biomedical Engineering, 38(2), pp.761-772.

[19] Bala, R., Sharma, A. and Goel, N., 2024. CTNet: convolutional transformer network for diabetic retinopathy classification. Neural Computing and Applications, 36(9), pp.4787-4809.

[20] Yu, M. and Wang, Y., 2022. Intelligent detection and applied research on diabetic retinopathy based on the residual attention network. International Journal of Imaging Systems and Technology, 32(5), pp.1789-1800.

[21] Sau, P.C. and Bansal, A., 2022. A novel diabetic retinopathy grading using modified deep neural network with segmentation of blood vessels and retinal abnormalities. Multimedia Tools and Applications, 81(27), pp.39605-39633.

[22] Saranya, P., Pranati, R. and Patro, S.S., 2023. Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models. Multimedia Tools and Applications, 82(25), pp.39327-39347.

[23] IDRID dataset link: https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid (Accessed on 02/05/2024).