Journal of Information Systems Engineering and Management
2025, 10(15s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Automatic Detection of Microaneurysms Using Modified UNET Architecture

T. Monisha Birlin[1*], C. Divya[2], J. John Livingston[3]
[1*]*Research Scholar, Center for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India, birlin.monisha87@gmail.com*
[2]*Assitant Professor, Center for Information Technology and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India, cdivyame@gmail.com*
[3]*Assitant Professor, Department of Computer Science and Engineering, Kamaraj College of Engineering and Technology,Virudhunagar, India, johnlivingstoncse@kamarajengg.edu.in*

ARTICLE INFO

Received: 28 Nov 2024

Revised: 22 Jan 2025

Accepted: 01 Feb 2025

ABSTRACT

Increasingly, more people throughout the world are suffering from Diabetic Retinopathy (DR). A high blood glucose level damages the retina, which is found at the back of the eye and causes vision loss, leading to DR. The earliest symptoms of DR are referred to as microaneurysms (MAs). Due to their small size, dusky color, and practically round shape, these MAs are easy for ophthalmologists to overlook during physical investigation. In this situation, reliable early MA diagnosis is useful to prevent DR before irreversible blindness. The manual detection of DR is a labor-intensive and time-consuming process. This manual detection leads to misclassification of DR images often. The proposed method uses modified UNET architecture to detect MAs in the IDRiD Dataset and yields an accuracy of 99.88%, a precision of 0.64768, an F1 score of 0.6765, and an AUC of 0.6765.

Keywords: Deep learning, Diabetic Retinopathy, Microaneurysms, UNET.

I. INTRODUCTION

DR causes permanent visual loss, and the number of people affected by this ailment is rising continuously [13]. A challenging topic in the world of computer vision is the early detection of DR. Clinical signs of DR, such as MAs, exudates, and hemorrhages, can be seen in retinal color fundus images. The identification of these symptoms is crucial for accurate diagnosis since they reveal the present condition of DR. According to Hann et al. [1] a common method for extracting these elements from fundus images is based on the morphology of digital images. Transparency and simplicity are the approach's benefits. Exudates on a fundus images are a crucial diabetic diagnostic indicator. To distinguish between normal and pathological retinal images, multiscale amplitude and frequency modulation techniques were the foundation of the approach put forth by Agurto et al. [2]. The modulations were applied to a collection of several tiny fundus image regions with various sorts of lesions. Then, using the amplitude-frequency response, a tiny region's feature vector was created. According to the scientists, there is a statistical distinction between diseased lesions and normal retinal structures based on amplitude and frequency modulation properties. Kazakh-British et al. [3] performed numerical tests using the following processing pipeline: first, blood vessels were extracted from fundus images using Frangi & Sato filters, and then lesions were detected using Convolutional Neural Network (CNN) classifier training. A fundus picture could be somewhere between 0.15 and 0.5MB in size. The DR deep learning method was put to the test by Gulshan et al. [4] using a dataset generated by a smartphone. This study demonstrates how a large audience can have access to diagnostic tools for DR. Mateen et al. [5] proposed a symmetrically optimized solution for region segmentation, feature extraction and selection, and fundus image classification using a visual geometry group network (VGGNet), Gaussian mixture model (GMM), Singular Value Decomposition (SVD), Softmax, and Principal Component Analysis (PCA). According to the authors, the VGG-19 model performed better in terms of computational efficiency and classification accuracy than AlexNet and the spatial invariant feature transform (SIFT). Guo et al. [6] developed a multi-lesion segmentation model. They developed the first miniature object segmentation network (L-Seg) that can divide all four kinds of lesions at once. Given that small lesion regions could not respond at a high degree of the network, they suggested a multi-scale feature fusion technique to overcome this problem. Taking into account the scenarios of both class-imbalance and loss-imbalance difficulties, they also recommended a multi-channel bin loss. Resizing, scaling, and data augmentation were the preprocessing techniques used. Tan et al. [7] proposed a 10-layer convolutional neural network to automatically, simultaneously separate and distinguish exudates, hemorrhages, and MAs. Cleopatra, Messidor, and DIARETDB1 are the datasets that were utilized. Normalizing the input image is the first step in segmentation. Mo et al. [8] published a unique method for locating exudates using cascaded deep residual networks. Both the E-Ophtha and HEI-MED databases, which are openly available, were used. No preprocessing or postprocessing methods were used by the authors. By fusing multi-scale shallow CNNs with performance integration, Chen et al. [14] used a shallow learning technique for the identification of DR. Base learners, also called as multi-scale shallow CNNs, are used for feature extraction. Performance integration was used to integrate the base learners' output. Here, performance integration was implemented in order to eliminate weak learners and boost the performance of strong base learners. Accuracy was the primary area of focus in this study. It should be enhanced by applying a useful segmentation technique. The key benefit of this strategy is the smaller training data set needed. Large retinal databases were trained to automatically extract features using deep learning techniques. A method for the automatic identification of DR using a capsule network was proposed by Kalyani et al. [15]. The basic capsule layer and convolution layer were used to extract the features. Class capsule and softmax layers were used for categorization. By eliminating the pooling layers, capsule networks were developed to avoid the difficulties in CNN. The concept of using watershed lines and ridge measurement with an SVM classifier to detect new vasculature in the optic disc was presented by Goatman et al. [16]. This technique facilitates quick feature extraction and classification while reducing manual labor. There are fewer instances where this technique fails to find the optic disc. Costa et al. [17] put his concept of detecting the existence of DR using incomplete data into practice. Instance encoding and image classification stages are jointly optimized in Multiple Instances Learning (MIL). In their investigation, lesions were used to distinguish between healthy and diseased images without knowing where the lesions were located. The key benefits of this strategy are that it is a weakly supervised method and does not utilize manually segmented training data. There is no requirement for preprocessing methods. There is no requirement for pixel data pertaining to various lesions. The severity of the cases was not graded in this study. The concept of using one-dimensional CNN on unrelated datasets was put up by Sun [18]. Batch Normalization (BN) and the CNN model were coupled to reduce dispersion gradients and boost accuracy. The training speed is also accelerated. Other irrelevant one-dimensional datasets can also use this technique. This strategy produced good results, however there were some issues because the statistical data collection produced inaccurate results. The proposed method uses slightly modified UNET architecture for the better performance in segmentation. Small lesions like MAs are difficult to assess using the symmetry-driven UNET model, and its conclusions are not more precise. Convolution and pooling layers make up downsampling, which functions as path reducing in the network and gathers global data. Convolutional and deconvolutional layers make up the up-sampling, which locates pixel points by acting as path expansion in the network [9]. The proposed architecture was trained and tested on IDRiD images and it produced promising results.

II MODIFIED UNET ARCHITECTURE

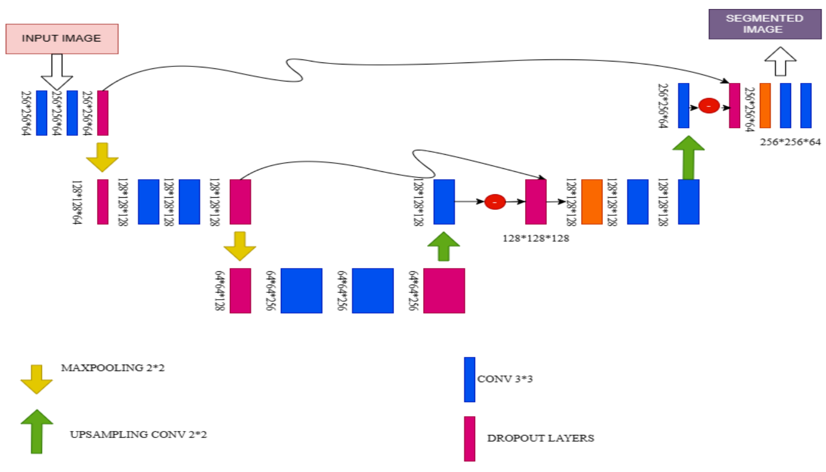

Fig. 1. Modified UNET Architecture

Fig 1. represents the modified UNET architecture with dropout layers. The dropout layers and the subtraction block are not present in the UNET architecture. Similar to UNET architecture, our proposed architecture contains an encoder and decoder structure. An encoder contains two downsampling operations, and the decoder consists of two upsampling operations. For object localization, upsampling techniques are used. Additionally, it increases resolution while decreasing depth. Downsampling methods are used for feature extraction. In addition, the depth increases. Encoders perform convolution operations (downsampling approach), and decoders do transpose convolutions (upsampling method). The skip connection's goal is to improve the information about microaneurysm that was probably lost through the downsampling process. The dropout layer is subtracted from the up sampled layer, and it is the novelty of the proposed architecture. The subtraction operation is shown in Fig. 1 using a red circle. The yellow straight arrow represents the downsampling operations along with the maxpooling of 2*2. The yellow straight arrow represents the upsampling operations. The blue color box represents the convolution operations of 3*3. The batch size is 4 and the number of epochs is 1000. The learning rate is $5e^{-5}$.

III. DATASET

The dataset we employed, IDRID [10], was derived from a patient's actual clinical fundus image acquired at an Indian ophthalmology clinic. Each image in the dataset was acquired using a Kowa VX-10 color fundus camera with a 50^{0} field of view that was positioned near macular region. Each image was produced in JPG format and has a resolution of 4288 x 2848. Out of 516 color fundus images, we chose 81 retinal images with MAs for our analysis. In accordance with the different lesion annotations, IDRID is split into a training set and a testing set.

IV. PREPROCESSING

Enhancing contrast between essential part of the image and its background is the main goal of preprocessing. The preprocessing methods we have included in our proposed method is Green Channel Extraction, Image Enhancement, and Batch Normalization. The green channel, out of three (red, green, and blue) channels, shows the most contrast between the MAs and the background. The size is reduced by eliminating the other two channels. Image enhancement is achieved through CLAHE. The performance of the image as a whole has been improved. The most effective method for locating MAs, according to our observation, was CLAHE [11]. CLAHE enhances the tiny features and provides better performance in locating MAs. In recent years, using CLAHE, high sensitivity, specificity, and AUC were obtained [12]. The normalization procedure involves transforming values into a predetermined format. Because of batch normalization, we are able to employ noticeably higher learning rates. Then the preprocessed images are segmented using the modified UNET segmentation algorithm. The proposed architecture is tested in the IDRiD Dataset, and it has produced promising results in terms of accuracy, F1 score, precision, and AUC.

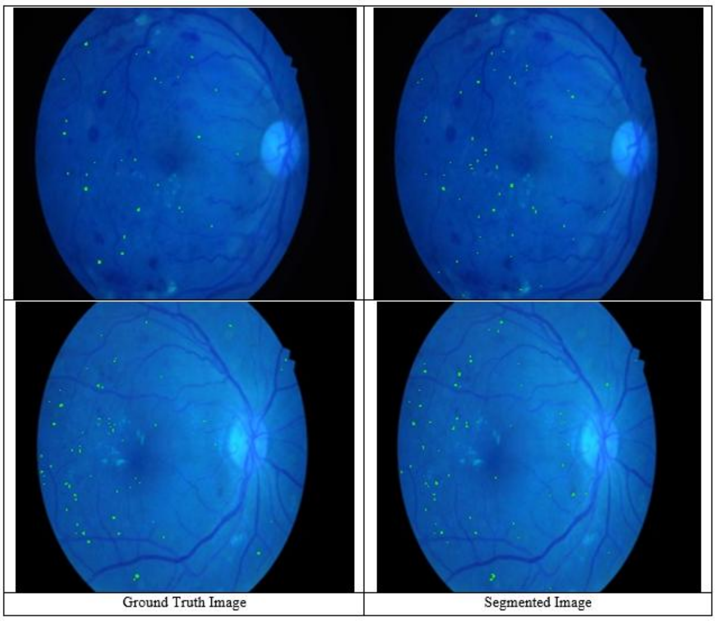

Fig. 2. Ground Truth Image versus Segmented Image for the proposed architecture.

Fig. 2 shows the segmented images and the ground truth images. The important parameters we have chosen for our segmentation algorithm are given in the following Table 1.

TABLE 1 LIST OF PARAMETERS

Name of the Parameter	***Values***
Learning Rate	$5e^{-5}$
Number of Epochs	1000
UNET Depth	5
Loss Function	5
Batch Size	4
Optimizer	Adam
Loss	Binary Cross Entropy (BCE)
Drop out	0.5%

V. LOSS FUNCTION

In our proposed architecture, BCE was used to identify whether MAs were present or not using probability values. The actual class output, which can only be either 0 or 1, is compared to each of the projected probabilities using BCE. Depending on how near or far the value is to the actual value, BCE is given in the following expression.

$$BCE = \frac{-1}{N}\sum_{k=1}^{N} \text{Yk} * \log(\text{p}(\text{Yk})) + (1 - \text{Yk}) * \log(1 - \text{p}(\text{Yk})) \qquad (1)$$

Equation (1) represents binary cross entropy loss function. Yk represents label. P(Yk) is the predicted probability for MAs and 1-P(Yk) represents the predicted probability for non-MAs. BCE is calculated for all the N points.

VI. RESULTS AND DISCUSSION

The proposed model is trained and tested using IDRiD and E-ophtha dataset. We got promising results after running 1000 epochs using PyTorch version 1.5.0+cu101. The following graph represents the performance measures in terms of accuracy, precision, F1 score, and AUC.

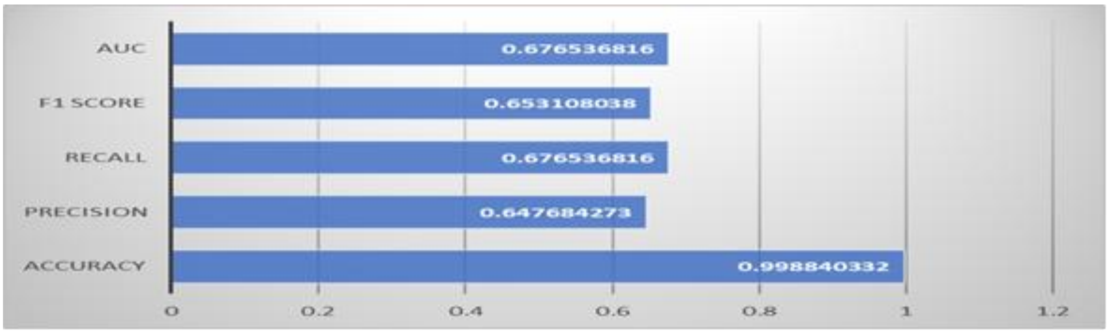

Fig. 3. The performance measures of the proposed model

It is clearly understood from the fig.3 that our proposed model achieved better results. Table 2 shows the values of achieved performance metrics in the IDRiD Dataset.

TABLE 2. PROPOSED MODEL PERFORMANCE

S.No	**Performance Metrics**	**Values**
1	Accuracy	0.9988
2	Precision	0.6477
3	Recall	0.6765
4	F1 score	0.6531
5	AUC	0.6765

TABLE 3. COMPARATIVE ANALYSIS OF ACCURACY WITH VARIOUS METHODS

S.No	Methods	Accuracy
1	Foo et al. [19]	0.7849
2	Guo et al. [6]	0.6852
3	Mo et.al. [8]	0.7962
4	Tan JH et al. [20]	0.8452
5	Xu et al. [21]	0.7867
6	Modified UNET	0.9988

Table 3 demonstrates that when compared to the majority of the existing approaches, our proposed method had higher accuracy.

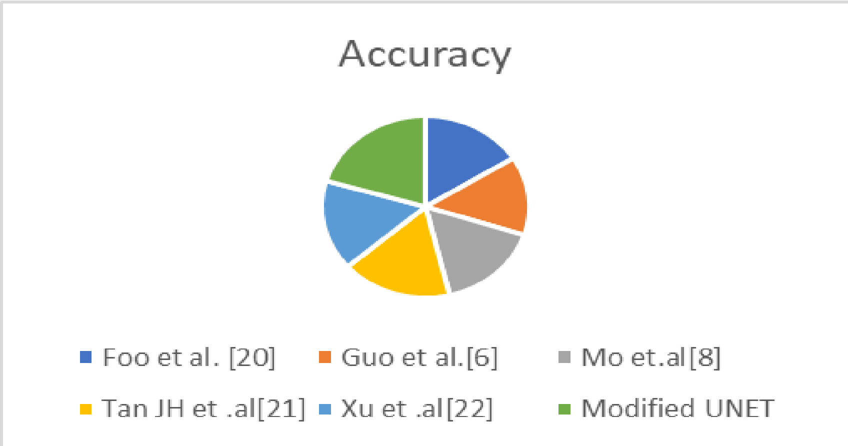

Fig.4. Comparison of Accuracy with Various Existing Methods

From the fig.4, it is clear that our suggested modified UNET architecture outperformed the majority of the already used segmentation techniques for the automatic identification of MAs

VII. CONCLUSION AND FUTURE ENHANCEMENT

Computerized diagnosis is highly sneeded because the number of diabetic patients is increasing every day. The number of ophthalmologists is very small when compared to the number of patients with DR. The manual detection of DR is also very expensive. We have tested our proposed model using the IDRiD dataset, which has an accuracy of 99.88%, a precision value of 0.6477, a recall of 0.6765, an F1 score of 0.6531, and an AUC of 0.6765. In the future, our modified UNET architecture is added with attention concepts to improve the performance further.

REFERENCES

[1] Hann, C. E., Revie, J.A., Hewett, D., Chase, J.G., Shaw, G.M. (2009). Screening for Diabetic Retinopathy Using Computer Vision and Physiological Markers. *Journal of Diabetes Science and Technology*, 3(4), 819–834.

[2] Agurto, C., Murray, V., Barriga, E., Murillo, S., Pattichis, M., Davis, H., Soliz, P., et al. (2010). Multiscale AM-FM Methods for Diabetic Retinopathy Lesion Detection. *IEEE Transactions on Medical Imaging*, 29(2), 502–512.

[3] Kazakh-British, N.S., Pak, A.A., Abdullina, D. (2018). Automatic Detection of Blood Vessels and Classification in Retinal Images for Diabetic Retinopathy Diagnosis with Application of Convolution Neural Network. *International Conference on Sensors, Signal and Image Processing*; ACM, 60–63.

[4] Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Kim, R., et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *Jama*; 316(22), 402–410.

[5] Mateen, M., Wen, J., Song, V., and Huang, Z. (2018). Fundus Image Classification Using VGG-19 Architecture with PCA and SVD. *Symmetry*, 11(1).

[6] Guo, S., Li, T., Kang, H., Li, N., Zhang, Y., Wang, K. (2019). L-Seg: An End-to-end Unified Framework for Multi-lesion Segmentation of Fundus Images. *Neurocomputing*, 349:52-63.

[7] Tan, M., and Le, Q.V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. International Conference on Machine Learning, 2019, arXiv: 1905.11946.

[8] Mo, J., Zhang, L., Feng, Y. (2018). Exudate-based Diabetic Macular Edema Recognition in Retinal Images Using Cascaded Deep Residual Networks. *Neurocomputing*; 290, 161-171.

[9] Ren, K., Chang, L., Wan, M., Gu, G., Chen, Q. (2022). An improved U-net Based Retinal Vessel Image Segmentation Method. *Heliyon*, 8(10), e11187.

[10] Porwal, P., Pachade, S., Kamble, R., et al. (2018). Indian Diabetic Retinopathy Image Dataset (IDRID): A Database for Diabetic Retinopathy Screening Research. *Data*, 3(3), 25.

[11] Pour, A.M., Seyedarabi, H., Jahromi, S.H.A., and Javadzadeh, A. (2020). Automatic Detection and Monitoring of Diabetic Retinopathy using Efficient Convolutional Neural Networks and Contrast Limited Adaptive Histogram Equalization. *IEEE Access*, 8, 136668 – 136673.

[12] Monisha, B.T., Divya, C., and Muthukumaran, N. (2022). Computerized Diagnosis of Diabetic Retinopathy based on Deep Learning Techniques. *International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 331-338.

[13] Monisha Birlin, T., and Divya, C. (2022). Comparison of Various Segmentation Techniques in Diabetic Retinopathy-A Review. *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough*, 73-79.

[14] Chen, W., Yang, B., and Li, J. (2020). An Approach to Detecting Diabetic Retinopathy Based on Integrated Shallow Convolutional Neural Networks. *IEEE Access*, 8, 178552 – 178562.

[15] Kalyani, G., Janakiramaiah, B., Karuna, A., et al. (2021). Diabetic Retinopathy Detection and Classification Using Capsule Networks. *Complex Intell. Syst*, 32, 707-721.

[16] Goatman, K.A., Fleming, A.D., Philip, S., Williams, G.J., Olson, J.A., Sharp, P.F. (2011). Detection of New Vessels on the Optic Disc Using Retinal Photographs. *IEEE Transactions on Medical Imaging*, 30(4), 972–979.

[17] Costa, P., Galdran, A., Smailagic, A., Campilho, A. (2018). A Weakly-Supervised Framework for Interpretable DR Detection on Retinal Images. *IEEE Access*, 6,18747 – 18758.

[18] Sun, Y. (2019). Neural Network of One- Dimensional Convolution Example of the Diagnosis of Diabetic Retinopathy. *IEEE Access*, 7, 69657 – 69666.

[19] Foo, A., Hsu, W., Lee, M., Lim, G., Wong, T.Y. (2020). Multi-Task Learning for Diabetic Retinopathy Grading and Lesion Segmentation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 13267 – 13272.

[20] Tan, J.H., Fujita, H., Sivaprasad, S., Bhandary, S.V., Rao, A.K., Chua, K.C, Acharya, U.R. (2017). Automated Segmentation of Exudates, Hemorrhages, Microaneurysms Using Single Convolutional Neural Network. *Information Sciences*, 420, 66–76.

[21] Xu, Y., Zhou, Z., Li, X., Zhang, N., Zhang, M., Wei, P. (2021). FFU-Net: Feature Fusion U-Net for Lesion Segmentation of Diabetic Retinopathy. *Biomed Res Int*.