

A Novel Deep Learning Approaches for Early Detection of DR and DME Using Fundus and OCT Images

Pranoti Nage¹, Charmi Chaniyara²

¹Thakur Shyamnarayan Engineering College, Information Technology Mumbai, India.

E-mail: pranotinage93@gmail.com

²St. Francis Institute of Technology, Computer Engineering Mumbai, Indi. E-mail: charmijaviya@gmail.com

Abstract

Diabetic patients suffered from various eye diseases such as Diabetic Macular Edema (DME) and Diabetic Retinopathy (DR) which causes optic disc damage and vision loss. Hence, early diagnosis of DR and DME is necessary for active treatment, but the slow movement of disease signs leads to difficulty in early detection. For early detection and severity grading of DR and DME, we proposed a novel deep learning approach. This research includes three major processes namely image quality assessment, bi-level segmentation, and multi-feature-based DR and DME detection. In first process, the image quality is estimated by generating threshold, if it has less quality then it performs preprocessing and data augmentation, otherwise it directly forward to the data augmentation process for increasing training efficiency. In preprocessing, we perform two process such as noise filtering and contrast enhancement. Noise removal is performed by Adaptive Threshold based Bilateral Filter (ATBF) which provides noise removed edge preserved image which increases the resolution of the image. For contrast enhancement, we proposed Honey Badger Optimization (HBO) which increases the image brightness by adjusting the intensity of the images. After completed image quality assessment, we perform image segmentation by using U-Net³⁺ which segments both optic disc and lesion by considering various features. The accuracy of segmentation is evaluated by Chi-square which reduces the false positive rate. Finally, feature extraction and DR and DME detection is performed by Attention based Hybrid SqueezeNet-VGG16 (AHS-VGG16) which extracts the features from segmented region and classified the images into three classes such as normal, DR and DME. The simulation of this research is conducted by Matlab simulation tool, and the performance of this research is evaluated based on different performance metrics by considering four types of datasets namely, Indian Diabetic Retinopathy Image Dataset (IDRID), Retinal Fundus Multi-Disease Image Dataset (RFMiD), Fundus and Fundus and OCT images, MESSIDOR

Keyword: Diabetic Retinopathy (DR), Diabetic Macular Edema (DME), Fundus and OCT images, Deep Learning, severity grading, optic disc and lesion segmentation

1. INTRODUCTION

Diabetic Mellitus is a chronic disease that causes various complications which affect health especially Diabetic retinopathy (DR). Approximately 381.9 million peoples are affected by diabetics. According to the statistics, there are approximately 90% of diabetic patients suffered from DR due to insufficient blood supply caused by damage of blood vessels [1]. The developed stage of DR leads to Diabetic Macular Edema (DME). It is caused due to fluid accumulation at macula because of retinal blood barrier disruption. Severe stage of DME leads to permanent damage (i.e., even blindness) [2]. Early detection of DR by the general clinical prognosis and emergency treatment is necessary to prevent the severe stage that reduces blindness. Ophthalmologic screening is performed by fundus images to detect DR which reduces the blindness risk by 50%. However, screening has a greater number of tests that increase the cost so, this is not possible in the present day. In addition, insufficient specialists in developing areas increase the risk factor. Automatic detection of DR by predicting the lesions from the retinal images with the help of image processing. Optical Coherence Tomography (OCT) images are used to detect DR and DME by retinal vessel imaging [3]-[5]. Various deep learning and machine learning classification algorithms are used to detect the DR and DME from the retinal images in an accurate manner but, the presence of noise reduces the detection accuracy. Fig 1 represents the DR and DME retinal images.

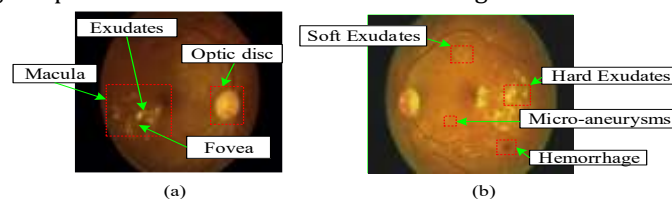


Fig.1 (a) DME image, (b) DR image

Several methods performed pre-processing by noise removal and contrast enhancement for image quality enhancement. Extraction of green channels out of the RGB color to enhance the contrast of the images to segment or highlight the lesions effectively. Various Regions of Interest (RoI) are combined using several segmentation techniques such as blood vessel segmentation, optic disc segmentation, and lesion segmentation by fuzzy system. Hemorrhages, soft exudates, hard exudates, microaneurysms lesions are segmented based on several properties such as color, size, etc. Microaneurysms appear as circular shapes with red color, hemorrhages are medium size and appear as red dots and exudates are white and yellowish colors. Some classifiers used deep learning approaches for performing supervised segmentation to increase the accuracy efficiently. Contrast limited adaptive histogram equalization (CLAHE) based segmentation is performed to increase the accurate detection of DR and DME. In many works, segmentation is performed to detect hard exudates and few works detect soft exudates (i.e., cotton wool spot) [6]-[9]. Feature extraction from the segmented images is performed by machine learning algorithms are used for feature extraction and classification such as Support Vector Machine (SVM), k-means clustering, and so on, but it takes much time for training larger dataset which leads to poor results in DR and DME detection [10]-[12]. Hence, deep learning algorithms are proposed for feature extraction and classification such as Convolutional Neural Network, Long-Short Term Memory (LSTM), VGG-16, and so on [13]-[17]. The features are extracted to analyze the morphological changes in the intercapillary areas and foveal avascular zone. Grading of DR is proposed in several works by extracting the features from the fundus images. The DR grading is mainly based on the lesions and it is detected by various pre-trained models. The DR grading includes five stages such as normal, mild, moderate, severe, Proliferated Diabetic Retinopathy (PDR), and Non-PDR. Likewise, DME is graded as normal or DME using ROIs of lesions [18]-[20]. Several datasets are used to train the deep learning model such as MESSIDOR, IDRiD, RFMiD, etc. In this research, we used various public datasets namely, IDRiD, RFMiD, fundus and OCT images, and MESSIDOR for detecting DR and DME.

A. Motivation & Objectives

The main aim of this research is to perform early detection of DR and DME in an accurate manner using fundus and OCT images with high accuracy and low false positive rate. This research addresses the problems of high false positive rate, high complexity and less detection accuracy. The major objective of this research is to detect DR and DME based on the severity using deep learning approaches. The other objectives of this research are listed as follows,

- To increase the quality of the input images by performing Image Quality Assessment that reduces the noise and increases the contrast of the image.
- To increase the detection accuracy by performing augmentation which train the images with multiple orientations.
- To achieve high accuracy by performing bi-level segmentation which segments the lesions and optic disc in the retinal images.
- To effectively classifies the images by performing multi feature based DR and DME detection based on low-level and high-level feature extraction, and classification.

The DR and DME detection had faced many challenges in recent years in terms of image pre-processing, feature extraction, and effective classification. The existing works provides achievable results but still lacks with effective solution we are motivated to solve the following problems.

- High false positive rate- The pre-processing of input images improves the overall performance. The existing works performs pre-processing however, noises are not pruned desirably which increases the false positive rate. In addition, segmentation process with improper feature consideration also increases the false positive rate.
- Consideration of single oriented images- Most of the existing works lacks with less detection accuracy, as they trained images with standard orientation however, lack of images with different orientations affects the detection accuracy.
- High complexity in feature extraction- Many of the existing works extract features by using convolutional neural networks which increases the complexity during feature extraction, as unwanted convolutional layers are formed during feature extraction thereby resulting in extraction of redundant features.

B. Research Contributions

This research mainly focuses on detecting DR and DME accurately using OCT and fundus images. The contributions of this research are listed as follows,

- In first, image quality is increased by performing image quality assessment, which is done by two performing two processes such as preprocessing, and augmentation. Preprocessing is achieved by performing noise removal using Adaptive Threshold-based Bilateral Filter (ATBF), and contrast enhancement using Honey Badger Optimization (HBO). Then, augmentation is performed by considering various angles.
- In second, bi-level segmentation is performed by using U-NET³⁺ algorithm which segments both optic disc and lesions in the retinal images which reduces the complexity of feature extraction and classification.
- Finally, we perform multi-feature-based detection by considering multi-features (i.e low-level and high-level) using Attention-based Hybrid SqueezeNet-VGG16 algorithm which perform feature extraction, feature fusion and classification for DR and DME detection which improves the accuracy of detection.

The performance of the proposed work is evaluated by various performance metrics such as, accuracy, precision, recall, f-measure, computation time, and ROC curve. The numerical results show that the proposed work achieved superior performance compared to existing works.

C. Paper Organization

The remaining sections of this paper is structured as follows, Section 2 illustrates the literature survey with research gaps, and Section 3 explains the problems existed in the previous works, Section 4 demonstrates the research methodology of the proposed work which includes algorithm procedure, pseudocode, and mathematical representations. Section 5 explains the experimental results which also includes dataset description, simulation setup, and comparative study. Section 6 concludes the research methodology with future directions.

2. LITERATURE SURVEY

In paper [21], proposed a novel method to detect DME using machine learning. Initially, color, spatial and textural features are extracted from the fundus images. Here, tree-based structure was proposed for selecting optimal features from the extracted features. Based on the features DME detection was performed by SVM algorithm. Here, severity grading was performed by measuring the distance between the fovea and optic disc and macula was located at the center. The performance of the proposed work was evaluated based on IDEDD, IDRiD, and DIARETDB1 datasets. Finally, the simulation result shows that the proposed work achieved better performance in DME detection. However, classification is performed by SVM algorithm which takes high much of time for training larger dataset that increases detection time which reduces the performance of the system. Authors in [22], proposed a novel algorithm to detect DME by using OCT images. In first, the OCT images were divided into background and RoI using k-means clustering. After that, segmentation was initiated to segment DME region using selective binary and Gaussian filtering regularized level set (SBGFRLS) algorithm, which was known as level set algorithm that was mostly used for segmentation of medical images. It includes three level set algorithms such as GAC, C-V and SBGFRLS. The experimental results demonstrate that the proposed work was achieved better performance in terms of sensitivity, specificity and precision. Here, DME detection is performed by SBGFRLS algorithm, however it segments the region from the raw input images that leads to high false positive rate due to the presence of noise and low contrast.

In paper [23], proposed a novel ensemble approach for detecting DR using deep learning. Initially, preprocessing was performed to improve the quality of the images which includes two process such as image resizing and augmentation. Here, CNN based ensemble learning namely densenet was performed for classifying the severity of DR which was based on the average values of the ensemble algorithms. The experimental result shows that the proposed work achieved better performance compared to other state of the art works. This paper proposed deep learning-based DR detection using fundus images [24]. Preprocessing was performed for the input images which includes RGB to gray conversion and image resizing. The entropy image and original images were trained in CNN, based on the training feature extraction and classification is performed. Then bichannel CNN was proposed for feature extraction and classification to detect DR. In this way, this research performed automatic DR detection using retinal images. The performance of the proposed work was evaluated based on sensitivity, specificity, accuracy, and ROC curve. However, feature extraction and classification is done by bichannel CNN which provides unwanted convolutional layers during feature extraction that leads to high complexity and less detection accuracy.

Authors in [25], proposed a deep learning framework for microaneurysm early detection. Initially, preprocessing is performed to improve the quality of the images. Then several features are extracted such as exudates, microaneurysms, and hemorrhages using DNN algorithm. Based on the features classification is performed based on CNN, VGG-16 and Inception-v3 algorithm. The performance of this work was evaluated based on two datasets such as DIARETDB1 and E-Ophtha which proved that the proposed work achieved better performance in classification accuracy. Here, limited features are considered for detecting microaneurysm which are not enough for accurate detection that reduces classification accuracy. Here, authors proposed an ensemble deep learning approach to detect DR using fundus images [26]. Initially, the fundus images are obtained from the Kaggle dataset and perform pre-processing by resizing and sampling (i.e., up sampling and down sampling) the images for dataset balancing. After perform sampling, rotate and flip the input image to increase the training time and reduce feature biasness. Several meta-algorithms are combined into a single predictive model to perform ensemble learning. There are five deep convolutional neural networks were combined such as, Inceptionv3, Resnet50, Dense 169, Xception, and Dense121. By this model the features of the input image were extracted and detect DR based on states such as normal, mild, moderate, severe, and proliferative. However, preprocessing is performed by considering the size and orientation of the image, but it is not sufficient to perform effective pre-processing that reduces the clarity of the image due to low contrast which reduces the accuracy. In paper [27], authors proposed an approach to detect microaneurysms from the retinal images in an automatic manner by supervised learning and radon transforms for DR screening. Initially, the input images were pre-processed by removing the background of the image with contrast enhancement and illumination equalization and performs masking to detect the landmarks of retinal vessels and optic nerve head by multi-overlapping windows and Radon transform. Finally, detection of microaneurysms was performed by the combination of supervised SVM classifier and radon transform. Three public datasets are used to test the proposed method. However, this

research used SVM classifier to detect microaneurysms but, it does not able to tolerate huge datasets that increases the training complexity which leads to high false positive rate. This paper proposed a deep learning method for automatic DR detection using fundus images [28]. Here, CNN was used for detecting DR by considering low- and high-level features. These features are extracted from the lesion in the images. From the extracted features, optimal features are selected based on PCA algorithm that reduced the complexity and overfitting. Based on the selected features classification was done by gradient boosting algorithm which classifies the DR into five classes. The performance of the proposed was evaluated based on MESSIDOR and EyePACS datasets. Here, PCA algorithm is used for optimal feature selection which reduces the complexity; however, it provides high information loss due to not select the principal components optimally. Table 1 illustrates the summary of literature survey.

Table.1 Summary of Literature Survey

References	Concept	Algorithm	Dataset	Limitations
[21]	Machine learning based DME detection and severity calculation	SVM	IDEDD, IDRiD, DIARETDBI	High detection time
[22]	DME detection using OCT images with the aid of better level set algorithm	k-means clustering, SBGFRSL,	OCT images	High false positive rate
[23]	Detection of Diabetic Retinopathy using ensemble learning algorithm	CNN based ensemble learning model	Kaggle dataset	Less detection accuracy
[24]	Diabetic Retinopathy detection using bichannel CNN	Bichannel CNN	Kaggle diabetic retinopathy	High complexity Less detection accuracy
[25]	Detection of microaneurysm using deep learning approach	CNN, VGG-16, Inception-v3	DIARETDBI and E-Ophtha	Less classification accuracy
[26]	Detection of Diabetic Retinopathy using deep learning method	Deep Convolutional Neural Network	Kaggle dataset	Poor image quality
[27]	Detection of microaneurysms in retinal images	SVM	MUMS-DB, DRIVE, MESSIDOR, DIARETDB1	High false positive rate
[28]	Diagnosis of diabetic retinopathy using CNN	CNN, PCA	MESSIDOR, EyePACS	High information loss
[29]	Detection and classification of diabetic retinopathy	k-means, SVM, binary tree	IDRiD	Less classification accuracy
[30]	Characterization of diabetic macular edema using OCT images	-	OCT images	Less true positive rate
[31]	Detection of diabetic retinopathy by performing lesion segmentation	EAD-Net	E-Ophtha, IDRiD	Poor image quality
[32]	Grading diabetic retinopathy using ELM algorithm	CNN and SVD, ELM	APTOS-2019, MESSIDOR	Poor results in grading DR
[33]	Early detection of retinopathy and severity grading	ResNet-Inception	IDRiD	High complexity in classification
[34]	Diagnosis of diabetic macular edema using ensemble learning approach	Ensemble of CNN	IDRiD, MESSIDOR	High false positive rate
[35]	Detection of diabetic retinopathy detection by extracting deep features	ResNet, RF	MESSIDOR, EyePACS	Less true positive rate

Machine learning based classification and detection of diabetic retinopathy was introduced [29]. This work contains image pre-processing, lesion-based feature extraction, and machine learning based classification. At pre-processing stage, the RGB image was converted into chromatic color, the converted images were fed for equalization and normalization. The contrast enhancement image was converted into RGB image. The enhanced image was provided for lesion-based feature extraction in which red and bright lesions features are extracted by threshold and morphological methods. Finally, the machine learning classifiers such as binary tree, k-means, and support vector machines are used for diabetic retinopathy detection and classification. However, extracted lesions features are fused and fed to machine learning classifiers such as binary tree, k-means, and support vector machines which increases the complexity during classification thereby leading to degraded classification accuracy. The optical coherence tomography images based diabetic macular edema detection was introduced [30]. This work employs three

processes namely image sampling, feature extraction, and classification. The sampling process was used for effective labelling of segmented parts in the images. The labelled parts are fed to feature extraction in which features with spatial, temporal, contextual information was extracted. The extracted features are fed to machine learning classifier for diabetic macular edema detection. Finally, based on the voting process the map was generated in order to find the characterization of diabetic macular edema. This work utilizes machine classifiers for final classification of diabetic macular edema detection however machine learning classifiers limits with less classification than deep learning classifiers. Neural network based diabetic retinopathy detection based on lesion segmentation was introduced [31]. This work employs two modules namely pre-processing module, and Encoder Attention Decoder module. Initially, the detection was improved by performing pre-processing steps such as augmentation and normalization. The pre-processed images were fed to Encoder Attention Decoder module, the encoder module employs down sampling process for feature extraction and creating corresponding feature map, the attention module gives importance to features with high spatial and context information by using two sub modules, and decoder module employs up sampling process for producing segmentation map from the attention features and giving segmented results. This work performs pre-processing of the input images such as augmentation and normalization however, the overall detection accuracy was degraded by illumination of the input images.

Machine learning based diabetic retinopathy grading based hybrid model was introduced [32]. This works employs phases such as pre-processing, feature extraction, feature selection, and classification. The ben graham and clache processes were used for pre-processing in which image normalization and pre-processing were taken place. The pre-processed images were fed to hybrid model of Convolutional Neural Network and Singular value decomposition network for feature extraction and feature reduction respectively. The reduced features are then fed to Extreme Learning machine classifier for final classification in which five diabetic retinopathy classes are classified. This work utilizes extreme machine learning for classification of diabetic retinopathy however, the extreme machine learning lacks with problem of overfitting which affects the accuracy of classification. The ResNet based early detection and grading of diabetic retinopathy lesions was introduced [33]. This work consists of three stages namely pre-processing, feature extraction, and severity classification. The input images are pre-processed for enhance the detection accuracy by using normalization and augmentation methods. The pre-processed images are provided to hybrid structure (ResNet-Inception) for feature extraction in which feature belonging to image vectors are extracted. Finally, fully connected layer contains classifies the diabetic retinopathy in pixel wise manner. Here, the extracted features are directly fed to the classifiers which increases the classification complexity thereby affecting classification accuracy.

The convolutional neural network based diabetic macular edema detection and diagnosis was introduced [34]. This work contains three processes namely pre-processing, diabetic macular edema detection, and grading based diagnosis. In pre-processing stage, the raw input funds images are fed for two sub processes namely quality enhancement and morphological operations for improve the image contrast. The pre-processed images were provided to cluster of CNNs which was arranged in hierarchical manner. The gated ensemble is provided in the CNN in order to accurate detection of diabetic macular edema. Based on the detection results grading of diabetic macular edema was performed. Here CNN clusters were used for learning the features from the pre-processed images which leads to formation of unwanted convolution layers leads to extraction of unwanted repeated features that increases the false positive rate. The detection of diabetic retinopathy was carried out the extracting deep features and machine learning based classification [35]. This work contains two phases deep feature extraction phase, and diabetic retinopathy detection phase. In deep feature extraction phase, Resnet_50 is utilized for extracting features from input images. The extracted deep features are created as feature vector map and fed to random forest classifier for detection of diabetic retinopathy based on grading. Based on the grading results, the severity of diabetic retinopathy can be classified as mild, moderate, and severe. This work direct extracts features from the raw input images however; the input images contain noises which must be pruned out that leads to less detection and classification accuracy respectively.

3. MAJOR PROBLEM STATEMENTS

The major research problems of the existing works are detailly explained in this section. The stacked autoencoders based diabetic macular edema diagnosis by extracting deep features was introduced [36]. Here, preprocessing, feature extraction, and classification were performed for detecting diabetic macular edema. The major problems of this research are listed as follows,

- This work considers Image Quality Assessment method as pre-processing process for improve the contrast of the image however, the noise in the input images were not considered which leads to less detection accuracy.

- This work utilizes four convolutional neural networks for deep feature extraction namely Squeezenet, resnet-50, inception-v3, and googlenet which increases the feature extraction complexity there results in decreased overall performance.
- Here, stacked autoencoders were used for integration and feature selection, and classification however the autoencoders are limits with high time consumption for training and information loss which affects the classification accuracy.

The detection of DR and DME was performed by using deep learning method from retinal images [37]. This research includes two processes such as localization and classification for detecting DR and DME. The research issues are explained as follows,

- Here, low-level features were considered for detecting DR and DME which were not enough for accurate detection, because it covers only partial information of the images that lead to less detection accuracy.
- All the features are extracted from the overall images instead of focusing on any affected region which leads to less true positive rate due to the complexity of extracting features from the entire images.
- Here, only single view-based training is performed for detecting DR and DME which does not provide an accurate result while the image is presented in different angle that reduces DR and DME classification accuracy.

The convolutional neural network based non proliferative detection of DR employing OCT images was introduced [38]. This work consists of processes such as pre-processing, extraction of features, and classification. The research problems are identified and defined as below,

- Here, the OCT images were pre-processed based on detection of fovea, segmentation of retina, extraction of patch, and arrangement of patch, however these processes include noise which affects the contrast of image. In addition, all the input images are pre-processed which consumes more time thereby resulting in unwanted errors.
- The input OCT images are pre-processed which enhances the accuracy of non-proliferative DR detection, however the images with different orientation (i.e., different views) were not focused which affects the overall performance of the system.
- The features extraction was done by many convolutional neural networks which extracts features for all the arranged patches separately which leads to high computational complexity and unwanted errors during classification.
- The extracted features by the convolutional neural networks were directly fed to the support vector machine classifier which was prone to classification errors as the classifier ability was not able to withstand with raw features which affects the DR detection accuracy.

This paper proposed a novel method for DR and DME detection using deep learning approaches [39]. This work includes two process such as disease detection and segmentation. The research problems are listed below,

- In this research, features were extracted by CNN for disease detection, however it generated multiple convolutional layers during feature extraction that leads to high complexity which increases false positive rate.
- Here, all the features were extracted from the raw input images for disease detection and segmentation which reduced detection and segmentation accuracy respectively due to the presence of noise and illumination in the images.
- Here, segmentation was performed based on fuzzy k-means clustering which localize the region of DR and DME, however it provided only approximate results not provide accurate segmentation results that leads to poor segmentation.
- Segmentation was done by considering optic disc which is not enough for detecting the disease in an accurate manner, however it leads to high false positive rate and low detection accuracy.

The early detection of macular region for diagnosis of Diabetic Macular Edema was introduced in this work [40]. The macular region was found by optical disc diameter, based on the optical disc diameter radius the severity of Diabetic Macular Edema was evaluated. The problems of this research are illustrated as follows,

- This work exploits fundus images for early detection of Diabetic Macular Edema based on macular center and region near macula however, the detection accuracy was affected with presence of illumination (i.e. low contrast image).
- Here, Diabetic Macular Edema was detected directly by two processes namely detection of macular center and region near macula respectively, while the entire process utilizes raw input images which increases the complexity during detection thereby leading to high false positive rate.

- This work utilizes support vector machine for classifying the Diabetic Macular Edema as mild, moderate, and severe however, the support vector machines limits with high training time which affects the classification accuracy.

Research Solutions: The above research problems are solved by proposed the following solutions. (1). The noise in the input images is reduced by pre-processing of images in which noise filtering (Adaptive Threshold-based Bilateral Filter) and contrast enhancement (Honey Badger Optimization) is performed for improve the detection accuracy. Data augmentation is performed for all the input images in which images are employed for flipping to degrees such as (45°, 90°,...etc) for improving the detection accuracy. The Image Quality Assessment is performed based on the threshold values in which two sub processes are taken place namely pre-processing and augmentation for reducing the noise, and enhance detection accuracy. (2). Bi-level segmentation is performed by using U-NET³⁺ algorithm in which both the lesions and optic disk diameter are considered. The Chi square method is used for estimating with ground truth images for assessing the similarity. Only the segmented part features are extracted from the bi-level segmentation process which increases the true positive rate and reduce the feature extraction complexity. (3). The feature extraction and classification is performed by attention based hybrid squeezeNet in which both the low and high level features are extracted, and only the important features with spatial and contextual information are considered for final classification by using attention mechanism. Both low level and high-level features are extracted for feature extraction by using hybrid squeezeNet-VGG 16 model for accurate detection for DR and DME. The Attention-based Hybrid SqueezeNet-VGG16 is used for classification of DR and DME which reduce the time for classification by considering only the features with high information from the attention layer.

4. RESEARCH METHODOLOGY

In this research, we mainly focus on accurate detection of DR and DME based on Fundus and OCT images. In addition, the classification of the severity level of the DR is performed using the extracted features and DME detection accurately. We have considered three datasets such as Indian Diabetic Retinopathy Image Dataset (IDRID), Retinal Fundus Multi-Disease Image Dataset (RFMiD), Fundus and OCT images and MESSIDOR. Fig 2 represents the architecture of the proposed work which includes major processes which are explained as follows,

- Image Quality Assessment
- Bi-Level Segmentation
- Multi-Feature based DR and DME Detection

In addition, it also includes the basic structure of the proposed algorithms such as UNet³⁺ and AHS-VGG-16 and sample outputs of DR and DME detection and its severity gradings.

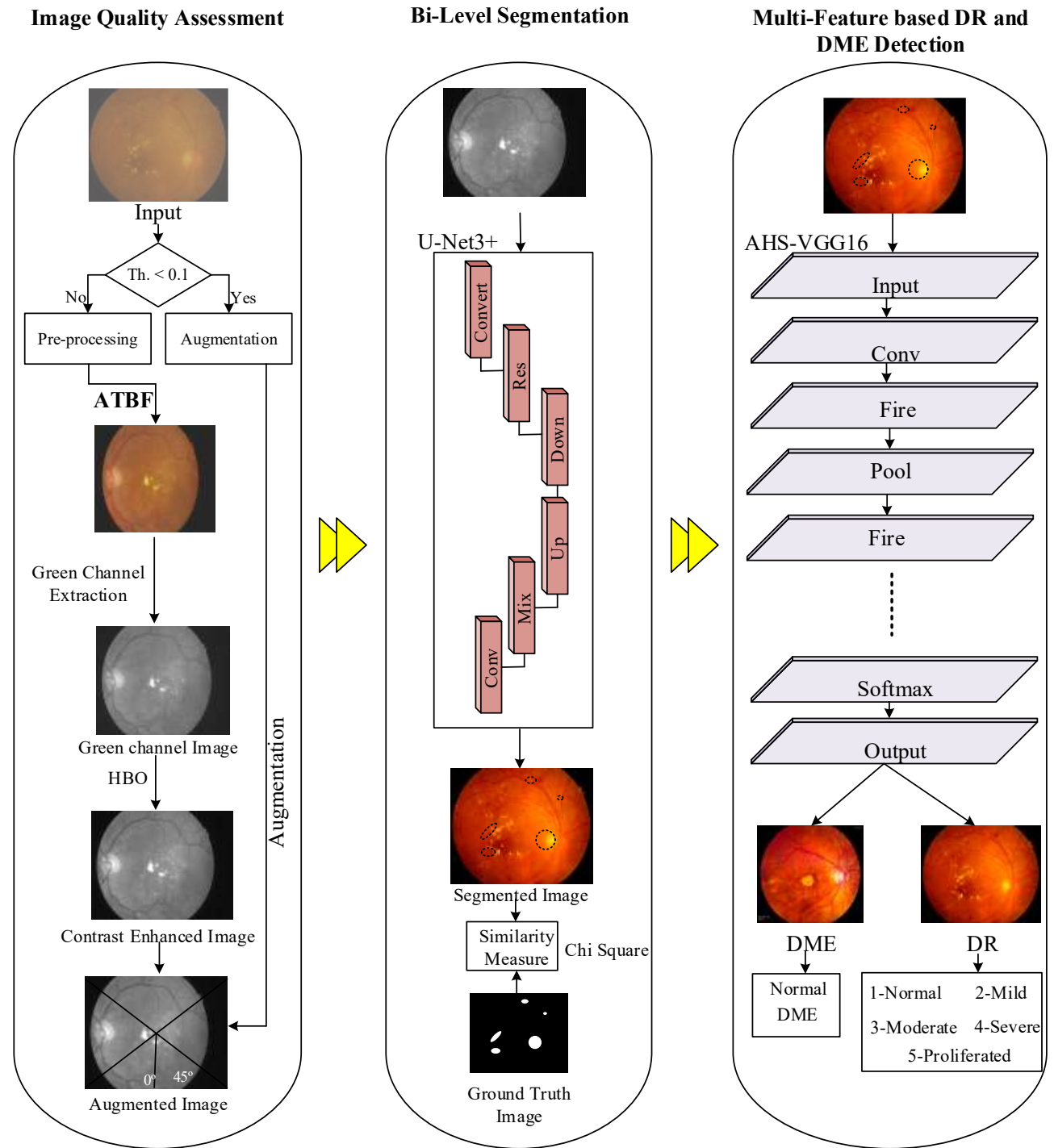


Fig.2 Architecture of the Proposed Research Methodology

4.1 Image Quality Assessment

Initially, the quality of all the input images is enhanced to detect the DR and DME diseases. In this process there are two sub-processes are implemented to increase the quality of the input images such as pre-processing and augmentation. The threshold value is generated based on the quality of images which is calculated as follows,

$$h(IQ) = - \sum_{i \in S} l(IQ) \log k(IQ) \quad (1)$$

where h represent the entropy function, and l , and k represent the probability distribution function, and IQ represent the image quality. If the threshold value is less than 0.1, directly goes to augmentation without performing preprocessing to reduce the complexity. Otherwise, perform preprocessing and augmentation. Two sub-processes are performed to enhance the quality of the input images such as pre-processing and augmentation. The detailed explanation of these two sub-processes are as follows,

a) Pre-processing

Pre-processing step is mainly implemented for improving the quality of the images by performing several processes such as noise filtering and contrast enhancement which are described as follows,

(i) Noise Filtering

Generally, input images have noise which degrades the detection accuracy. Noise filtering is performed to remove these noises from the respective input images using the Adaptive Threshold-based Bilateral Filter (ATBF) approach. Here, we adaptively changed the weight values of the photometric similarity to reduce the noise from the image. The adaptive threshold is generated based on sample trimming. The range of sample trimming is defined as follows,

$$f_{i,j} \in \varphi, s. |f_{i,j} - w_\mu| \leq \delta \cdot w_\rho \quad (2)$$

where $f_{i,j}$ represent the sample intensity of the reference window, δ represent the depth of trimming, w_μ represent the mean of the samples in the window, and w_ρ represent the standard deviation in the window which are expressed as follows,

$$w_\mu = \frac{1}{X^2} \sum_{y=1}^x \sum_{x=1}^x f_{i,j} \quad (3)$$

$$w_\rho = \sqrt{\frac{1}{X^2} \sum_{y=1}^x \sum_{x=1}^x (f_{i,j} - w_\mu)^2} \quad (4)$$

where X represent the reference window size, if eqn (2) is correct, then the window reserved for bilateral filtering, if it is wrong, then it will be eliminated. After calculated adaptive threshold, the bilateral filtering weight values are estimated as follows,

$$Bf(n) = c^{-1}(n) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\epsilon) \cdot g(\epsilon, n) \cdot S(p(\epsilon), p(n)) d\epsilon \quad (5)$$

where $g(\epsilon, n)$ represent the geometric closeness among PUF of n and pixel of surroundings ϵ , $(p(\epsilon), p(n))$ represent the photometric similarity among pixels of surroundings and PUF of n , and $c^{-1}(n)$ represent the coefficient of normalization which is defined as follows,

$$c(n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\epsilon, n) S(p(\epsilon), p(n)) d\epsilon \quad (6)$$

Bilateral filtering includes two weight functions namely photometric similarity and geometric closeness. The weight value of the geometric closeness is calculated as follows,

$$g(\epsilon, n) = e^{-1/2 \left(\frac{d(\epsilon, n)}{\gamma_d} \right)^2} \quad (7)$$

where $d(\epsilon, n)$ represent the distance between surrounding pixels (ϵ) and central pixel (n), and γ_d represent the diffusion factor of geometric function. The photometric similarity of the window is calculated as follows,

$$S(p(\epsilon), p(n)) = e^{-1/2 \left(\frac{p(\epsilon) - p(n)}{\beta_d} \right)^2} \quad (8)$$

where β_d represent the diffusion factor of photometric similarity. To achieve good results by selecting β_d and γ_d selected optimally. After that, the samples of ϵ are substituted with the trimmed samples of φ . Here, trim depth value is fixed ($\delta = 1.0$) and the outside ranges are neglected. This filter removes the noise, unwanted texture and preserves the edge of the images based on average pixel intensity to smoothen the image by combined weight values of the similarities. In addition, the threshold computation is adaptively performed to remove the noise based on the resolution of the input images.

(ii) Contrast Enhancement

After removing noise from the images, the green channel of the noiseless image is extracted from the RGB colors which improve the intensity distribution and luminance of the image that increases the accuracy of DR and DME detection. Generally, retinal images have low contrast, but the microaneurysms are clearly presented in green channel because of its high contrast. Hence, we perform green channel extraction before executing contrast enhancement. To enhance the contrast of the image, the information of red and blue channels in the images are removed. The contrast of the green channel extracted image is enhanced using Honey Badger Optimization (HBO) algorithm by optimizing the pixel values to improve the brightness of the image. In first, the position of the honey badger is defined as follows,

$$h = \begin{bmatrix} h_{11} & h_{11} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & \dots & \vdots \\ h_{m1} & h_{m2} & \dots & h_{mn} \end{bmatrix} \quad (9)$$

$$h_i = [h_i^1, h_i^2, \dots, h_i^d] \quad (10)$$

where h_i represent the i th position of the honey badger. The proposed honey badger includes four major processes such as initialization, searching, updation, and output generation. Initially, the honey badger is initialized with respect to number of solutions that is defined as follows,

$$h_i = L_{bi} + \text{Ran1} \times (U_{bi} - L_{bi}) \quad (11)$$

where, L_{bi} represent lower bound and U_{bi} represent upper bound of search space, h_i is current position of the honey badger, Ran1 represent the random number between 0-1. The smell intensity is related to prey concentration strength, if the intensity is high then the motion also fast which is calculated as follows,

$$S_{li} = \text{Ran2} \times \frac{\zeta}{4\pi D_i^2} \quad (12)$$

$$\zeta = (h_i - h_{i+1})^2 \quad (13)$$

$$D_i = h_{\text{prey}} - h_i \quad (14)$$

where S_{li} represent the prey smell intensity, and ζ represent the concentration strength which denotes the prey location, D is the distance between badger and prey, and the values of Ran2 is between 0 to 1. Then update the density factor (ϑ) of HBO which is used to confirm the transition between exploration and exploitation. The updation of density factor is defined as follows,

$$\vartheta = \partial \times \exp\left(\frac{-\hat{1}}{\hat{1}_{\text{Max}}}\right) \quad (15)$$

where ∂ represent the constant which value is greater than or equal to one, the default value of constant is two, and $\hat{1}_{\text{Max}}$ represent the count of maximum iterations. Finally, the output phase that is updation of agent position is divided into two stages such as digging stage and honey stage. In digging stage, the new position of honey badger is formulated as below,

$$h_{\text{new}} = h_{\text{prey}} + f \times \vartheta \times S_i \times h_{\text{prey}} + f \times \text{Ran3} \times \vartheta \times D_i \times [\cos(2\pi \text{Ran4}) \times [1 - \cos(2\pi \text{Ran5})]] \quad (16)$$

where h_{new} represent the honey badger current position, f represent the coefficient of flag which can change the direction of search space, and ϑ represent the capability of honey badger to get food. The mathematical formulation of search direction is defined as follows,

$$f = \begin{cases} 1 & \text{if } \text{Ran6} \leq 0.5 \\ -1 & \text{else} \end{cases} \quad (17)$$

The optimal solution of HBO is formulated as follows,

$$h_{\text{new}} = h_{\text{prey}} + f \times \text{Ran7} \times \vartheta \times D_i \quad (18)$$

In this way, contrast enhancement is performed in this research, here the intensity of the pixel is adjusted or changed optimally based on the current position of the honey badger. Implementation of this algorithm is mainly based on its search space, high convergence speed, and balancing exploration-exploitation.

b) Augmentation

The accuracy of the DR and DME detection is improved by performing image augmentation considering rotating, translation, scaling, and flipping to alter the pre-processed images to improve the size of the datasets that enhance the training and testing performing in detection of DR and DME. Fig 3 represents the sample of data augmentation.

(i) Flipping: It generates the mirror image of the retinal image with respect to both vertical and horizontal format. The retinal images have both left and right-side images, we make the uniform direction of the retinal images for reducing complexity during disease detection.

(ii) Translation: It translate the whole retinal images with respect to translation vector based on precise direction. It provides more geographical information of the images rather than focusing on single location.

(iii) Rotation: It rotates the retinal image into both clockwise and anticlockwise between the range of $0^\circ - 359^\circ$ which determined the augmentation safety. In addition, retinal images classified into four types of rim thickness based on the degree which are Inferior (I), Superior (S), Nasal (N), and Temporal (T). The rim thickness is helps to segment the optical disk from the retinal image. The determination of rim thickness is calculated as follows,

$$\text{I} = 0^\circ \leq 45^\circ \text{ and } 316^\circ \leq 359^\circ \quad (19)$$

$$\text{S} = 46^\circ \leq 135^\circ \quad (20)$$

$$\text{N} = 136^\circ \leq 225^\circ \quad (21)$$

$$\text{T} = 226^\circ \leq 315^\circ$$

(22)

(iv) Scaling: It helps to resize the images into uniform size within a dimension range. The scaled images are used for training which reduces the complexity of managing the images with various sizes.

The training is based on multi-view of the images which increases the detection rate and accuracy of segmentation and DR and DME detection.

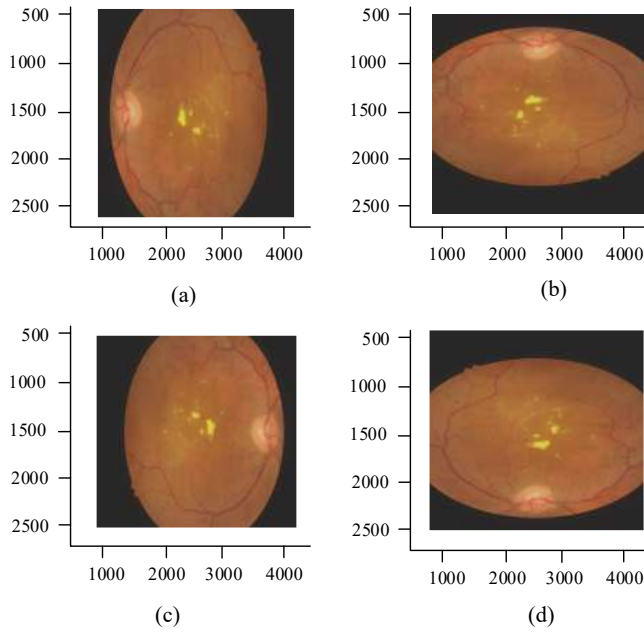


Fig.3 Sample images for Data Augmentation (a) sample input image (b) Rotation on 90-degree (c) Rotation of 180-degree (d) Rotation of 270-degree

Pseudocode for Preprocessing

Input: Fundus and OCT images (I)

Output: Preprocessed image

For each I do

 Verify IQ based on threshold using eqn (1)

If (IQ < 0.1)

 Initialize augmentation ()

 Compute flipping

 Compute translation

 Compute Rotation

 Compute scaling

Else

 Perform noise removal using eqn from (2) to (8)

 Extract green channel component

 Execute contrast enhancement using HBO

 Initialize population of HB using eqn (9)

While $i < I_{\max}$ **do**

 Density factor updation using eqn (15)

For $i = 1$ to HB_i **do**

 Calculate S_{li} using eqn (12)

If $Ran6 \leq 0.5$ **then**

 Update the current position from digging stage using eqn (16)

Else

 Update current position from honey stage using eqn (18)

End if

End for

End while

 Go to step

 Return preprocessed image

End for

4.2 Bi-Level Segmentation

For effective feature extraction with low complexity, bi-level segmentation of the augmented image is performed in this research such as lesion segmentation and optic-disc segmentation using U-NET³⁺ algorithm. Initially, lesion segmentation is performed by considering several features such as color, size, intensity, shape, and position to segment various lesions such as hemorrhage, soft exudates, hard exudates,

and microaneurysms which are shown in table 2. Optic disc segmentation is performed by considering the diameter of the optic disc and additional information of rim thickness ($\ell, \tilde{N}, z, T$) which reduces the feature extraction complexity and improves the detection accuracy of DR and DME. Fig 4 represents the structure of U-NET³⁺.

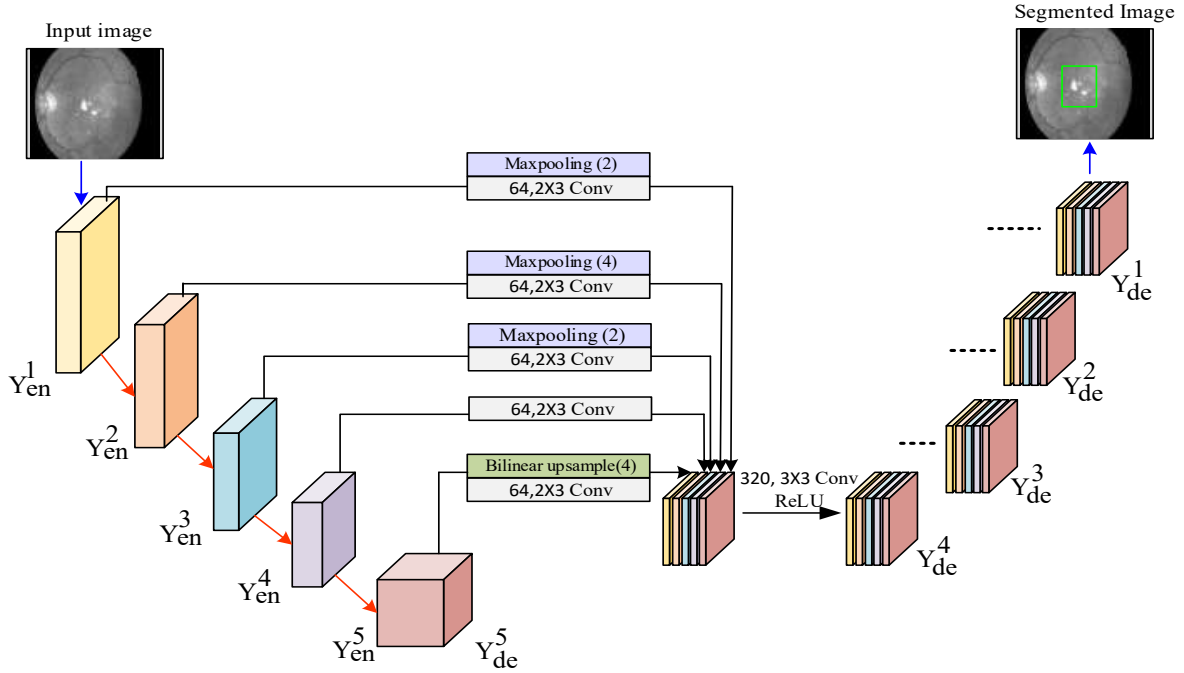


Fig.4 Architecture of Proposed U-NET³⁺

The proposed U-NET³⁺ algorithm also removes the semantic ambiguity during skip connection and used the benefits of full-scale skip connections. Compared to UNet, and UNet++, the proposed **U-NET³⁺** associates the multi scale features by modify or redesign the skip connections for providing accurate results. Every decoder layer of U-NET³⁺ generates both same scale feature map and smaller feature map, and decoder generates larger scale feature map which are helps to capture both coarse grain and fine grain information in full scales. The feature maps from encoder layers are directly established in the decoder. The intra decoder skip connection chain helps to transmit the segmentation information using bilinear interpolation from large scale decoder. To merge the information with semantic information, feature aggregation is performed to generate concatenate feature map which includes 320 filters with the size of 3×3 , and ReLU activation function. The skip connection of the proposed U-NET³⁺ is formulated as follows,

$$Y_{de}^i = \begin{cases} Y_{en}^i & i = n \\ g \left(\left((\mu(\emptyset(Y_{en}^j)))_{j=1}^{i-1}, \mu(Y_{en}^i), \mu(\alpha(Y_{de}^j))_{j=i+1}^n \right) \right), & i = 1, \dots, n-1 \end{cases} \quad (23)$$

where μ represent the convolution operation, and g represent the mechanism of feature aggregation, $\emptyset$ indicates the upsampling operation and α represent down sampling operation. The decoder parameters are calculated as follows, $\mathcal{Q}_{\alpha-de}^i = \mu_k \times \mu_k \times \left[\left(\sum_{j=1}^i D(Y_{de}^{j+1}) + D(Y_{de}^j) + \sum_{j=i+1}^n D(Y_{de}^j) \right) \times 64 + D(Y_{de}^j)^2 \right]$ (24)

where $\mathcal{Q}$ represent decoder parameters, and μ_k is a kernel size of convolutional layer, and D represent the node depths. The full-scale aggregated features maps adopted with deep supervision method to learn the aggregated features. The realization of deep supervision is performed to fed the decoder stage into convolutional layers afterwards bilinear upsampling and sigmoid function. It provides the segmented region of the fundus and OCT images. To enhance the accuracy and similarity of the segmentation results, we compared the segmentation image to ground truth image using Chi-Square (χ^2) algorithm. The similarity calculation is formulated as follows,

$$\chi^2 = \sum_{i=1}^n \frac{(\eta_i - \beta_i)^2}{\beta_i} \quad (25)$$

where η_i represented the segmented output region, β_i represent the expected output region that is ground truth region which increases segmentation accuracy and reduce false positive rate.

Table.2 Features and Lesions of DR and DME

Features	Lesions	Description
----------	---------	-------------

Color	Hemorrhage	It seems like a bigger spot in the retina, the size of hemorrhage is larger than $125\ \mu\text{m}$
Size		
Intensity		
Shape	soft exudates	It seems like a white spot which is because of nerve fiber swelling.
Position	hard exudates	It seems like a yellow spot in the retina which is because of plasma leakage.
	microaneurysms	It is an early symptom of DR which seems like a red dot in the retina because of vessel walls weakness. The size of microaneurysms is smaller than $125\ \mu\text{m}$

4.3 Multi-Feature based DR and DME Detection

The feature extraction is performed from the segmented images based on two level features such as high level and low-level features. In this research, spatial ($\mathcal{h}$), structural ($\mathcal{j}$), and statistical ($\mathcal{k}$) features are considered as high-level features and color ($\mathcal{c}$), shape ($\mathcal{s}$), and texture ($\mathcal{t}$) are considered as low-level features which are extracted from the segmented region. After feature extraction, feature fusion is performed to combine both the high-level and low-level features for DR and DME classification. DR is classified based on these fused features and the classified results regarding the severity grades which are based on the severity of DR. The proposed AHS-VGG16 deep learning architecture is used to perform both feature extraction and classification.

The proposed AHS-VGG-16 includes 12 fire layers and 4 convolutional layers in which 3 convolutional layers are defined as fully connected layers of VGG-16 which are shown in fig 5. Then, we incorporate scale layers to the first convolutional layer and all fire layers. Then, we substitute the fire layer to the 2nd convolutional layer of VGG-16 architecture. The input layers have 3×3 size of filters. To estimate the fire module parameters, the input channels are multiplied by the filters. Every max pooling layer is designed by considering 3×3 kernel size following by activation maps. Here, convolutional layers generate the activate map with space resolutions. The activation map wight and height size are determined based on two features such as input size and layer choice. In addition, the proposed AHS-VGG16 architecture includes convolutional layer with 1×1 filter size and it is forward to the expand layer with 1×1 and 3×3 convolutional filters.

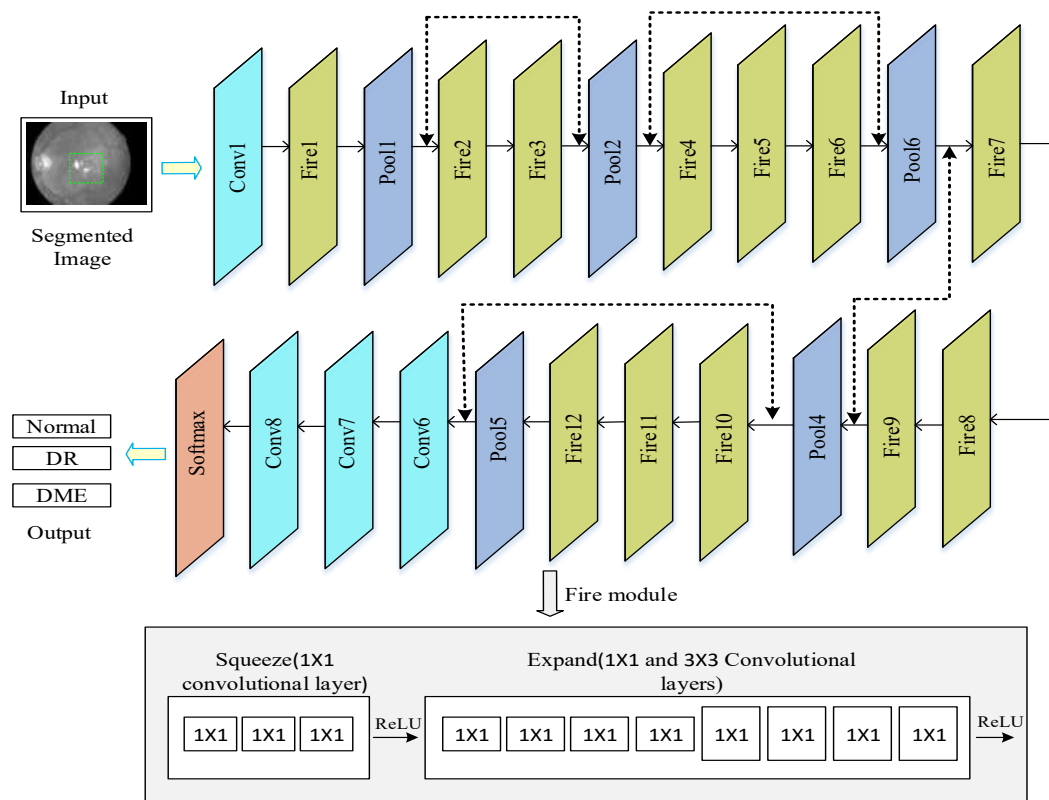


Fig.5 Architecture of Proposed AHS-VGG16

The shortcut connection of AHS-VGG16 is formulated as follows,

$$O = f(z, \{w_i\}) + z \quad (26)$$

where O represent the output of shortcut connections, f is a fire layer, z represent the output of pooling layer. The shortcut connections are walk across three fire layers. The result of shortcut connection is involved to the pool2 and fire6 output. Here, fire 3 layer has an output channel with the size of 128 while fire6 layer has a output layer with the size of 256. The pool2 is included the new convolution layers (256 output channel). The connection connects the fire9 and pool3 outputs, where the fire9 has 512 output channel and fire6 has output channel with the size of 256. Hence, the new convolutional layer (256) is including with the pool3. Finally, the connections fused the pool4 and fire12 output, since the fire 12 and fire9 has same size of output channel (256). Therefore, it does not any new convolutional layers between fire12 and fire9 which are previously equalized with 512 channels. The final result of convolutional layers is computed as follows,

$$cl = \sigma(f^{3 \times 3}[M_p]) \quad (27)$$

where σ represent the sigmoid function, and $M_p \in r^{h \times w}$ represent the width and height of the pooling layer. The convolutional layer extract both low level and high-level features which are fused in the attention layer which selects only limited featured thereby reducing unwanted and redundant features. The extracted features are classified by softmax layer which classified the images into three classes such as normal, DR, and DME. The output of classification is defined as follows,

$$O\left(p = \frac{s}{t}\right) = \frac{e^b}{\sum_i e^{b_i}} \quad (28)$$

where s and t represent the probability which are get from the softmax layer. For achieving better accuracy, we minimize the loss function using binary cross entropy which are computed as follows,

$$L = -\frac{1}{Q} \sum_{q=1}^Q \left[T_q \times \log(w_o(I_q)) + (1 - T_q) \times \log(1 - w_o(I_q)) \right] \quad (29)$$

where Q represent the count of training samples, T_q represent the target for training samples q , I_q represent the input of training samples q , and w_o weight values of the layers.

Table 3 represent the layers present in the proposed AHS-VGG16 which also shows the dimensions of fire layers such as $\$1 \times 1$, $E1 \times 1$ and $E3 \times 3$ where $\$1 \times 1$ represent the number of filters (1×1) present in the squeeze layer, and $E1 \times 1$ count of filters (1×1) present in expand layer, and $E3 \times 3$ represent the number of filters (3×3) present tin the expand layer. We defined $\$1 \times 1$ is always less than $E1 \times 1 + E3 \times 3$ for limiting the number of input channels for the filters.

Table.3 Layers of Proposed AHS-VGG16 Model

Layers	$E3 \times 3$ (expand)	$E1 \times 1$ (expand)	$\$1 \times 1$ (squeeze)
Fire 1	32	32	8
Fire 2	64	64	16
Fire 3	64	64	16
Fire 4	128	128	32
Fire 5	128	128	32
Fire 6	128	128	32
Fire 7	256	256	64
Fire 8	256	256	64
Fire 9	256	256	64
Fire 10	256	256	64
Fire 11	256	256	64
Fire 12	256	256	64

After detecting DR and DME, we estimate the severity level of these disease. The severity of DR depends on the location (β), number of lesions (ω) and thickness of exudates (υ). The severity ($\hat{\$}$) is calculated by using threshold which is generated by entropy functions. The threshold generation is formulated as follows,

$$h(\hat{\$}) = - \sum_{\hat{\$} \in \hat{\$}} l(M) \log k(M) \quad (30)$$

$$M = \sum m1 + m2 + m3 \quad (31)$$

$$m1 = l(\beta) \log k(\beta) \quad (32)$$

$$m2 = l(\omega) \log k(\omega) \quad (33)$$

$$m1 = l(\upsilon) \log k(\upsilon) \quad (34)$$

where M represent the metrics which are taken for generating threshold value, based on threshold the severity level is classified into five grades, the value of threshold between 0 to 1. The severity level is computed as follows,

$$G1 = 0.1 \geq h < 0.25 \quad (35)$$

$$G2 = 0.25 \geq h < 0.5 \quad (36)$$

$$G3 = 0.5 \geq h < 0.75 \quad (37)$$

$$G4 = 0.75 \geq h \leq 1 \quad (38)$$

The severity grades range from 1 to 4 in which grade 1 denotes as mild DR, grade 2 represents moderate DR, grade 3 denotes severe DR and grade 4 denotes proliferated. The gradings are shown in table 4. After DR detection, DME classification is performed based on the fused features and the severity grades of DR. The classification of DME results in normal or DME. In this way, we can detect DR and DME in early stage by considering the important features.

Table.4 Severity Gradings of DR

Grades	Severity	Characteristics
1	Mild	Only presence of microaneurysms
2	Moderate	Presence of intraretinal hemorrhages and microaneurysms, hard exudates
3	Severe	Microaneurysms present in all four quadrants, venous beading in two or more quadrants
4	Proliferated DR (PDR)	Neovascularization of retina, optic disc and iris, and vitreous hemorrhage presence

Pseudocode for DR and DME detection

Input : Segmented image
Output: Severity grading of DR and DME
Begin
 Initialize features (lf, hf)
 $lf \leftarrow \{c, s, t\}$
 $hf \leftarrow \{h, j, k\}$
 Initialize AHS-VGG16
 For $i \leftarrow 0$ to n **do**
 Extract lf and hf from segmented image
 Substitute fire layer to second convolutional layer
 Design a pooling layer with 3×3 kernel size
 Execute shortcut connections using eqn (26)
 Perform shortcut connection fusion
 Extract the features by convolutional layers M_p
 Execute classification by softmax layer using eqn (28)
 Compute loss function using eqn (29)
 Obtain classification results (N, DR, DME)
 End for
 Compute threshold value using eqn (30) for severity grading
 Obtain severity grading from eqn (35) to (38)
End

5. EXPERIMENTAL RESULTS

This section explains the experimental results of the proposed work. This section includes four sub sections such as dataset description

A. Dataset Description

This section describes the dataset which are considered in this research for detecting DR and DME. here, we used four types of datasets which are explained as follows,

IDRiD: This dataset includes retinal images with Indian population which is the first dataset demonstrative Indian population. This dataset includes DR and DME images and its severity. It helps to analysis the early signs of DR detection. It includes 516 images in which 413 are used for training and remaining are used to testing.

RMFiD: This dataset has 3200 fundus images which are captured by various fundus camera with forty-six conditions. The dataset classified into three sub sections such as training, testing and evaluation. Two categories of annotations are present in this dataset such as retinal image screening, and retinal image classification.

Fundus and OCT images: This dataset includes 100 OCT images with normal DR and DME categories. Here, 10 images are represented as normal images, others are DR and DME classification. 90 patients are affected by DR or DME where all patients are scanned by the same devices by the same medical practitioner to reduce potential methodical unfairness.

MESSIDOR: This dataset includes 400 retinal images. here, grade of DR is provided by macular edema risk, all the images are captured by 3CCD camera. It includes five stages of DR such as normal, mild, moderate, severe, and PDR, and two stages DME such as normal, and DME.

B. Simulation Setup

The simulation setup of the proposed model is conducted by Matlab R2020a which helps to improve the performance of simulation results. Table 5 illustrates the specification of both hardware and software. The simulation of this research is performed the processes of image quality assessment, bi-level segmentation, and multi-feature-based DR and DME detection. Fig 6 represents the training and testing of AHS-VGG-16.

Table.5 System Specifications

Hardware Configurations	RAM	6GB
	CPU	3.00GHz
	Hard disk storage	1 TB
	MATLAB	R2020a
Software Configurations	Processor	Intel (R) core (TM) i5-4590S
	Operating System	Windows 10

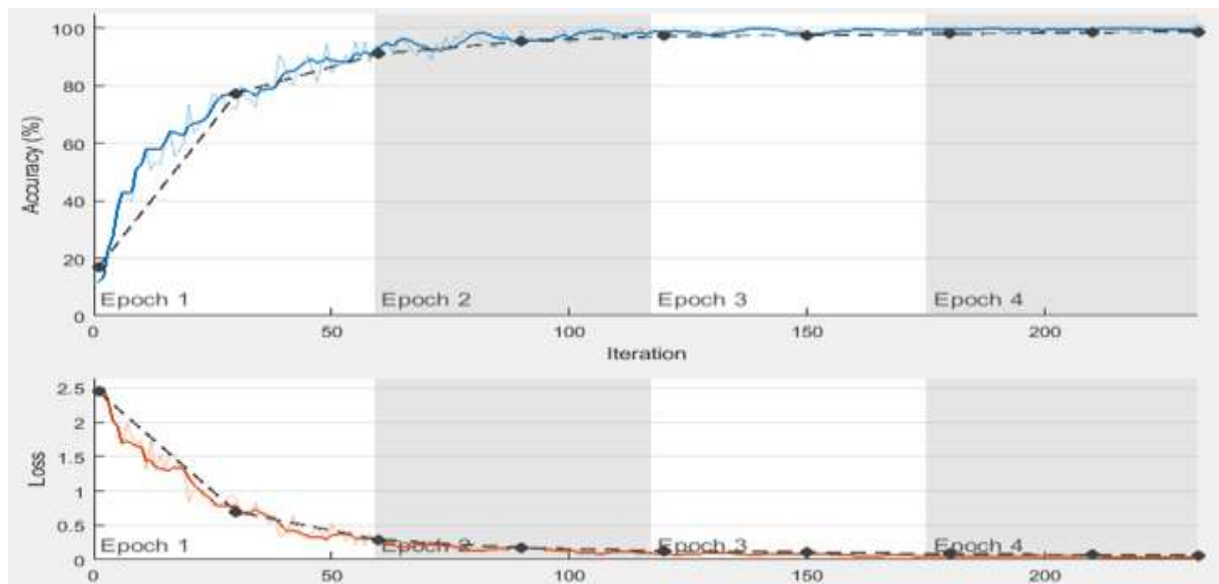


Fig.6 Accuracy and Loss of Proposed AHS-VGG16

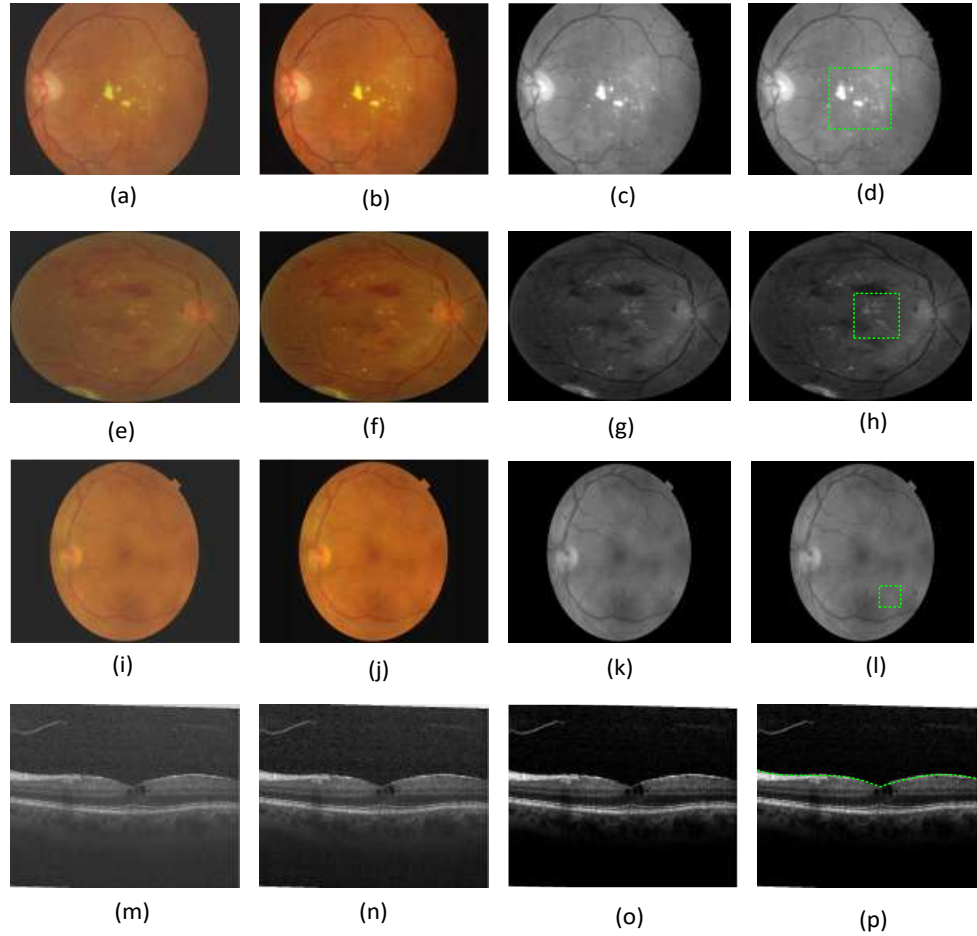


Fig.7 Simulation Results

Fig.7 represents the simulation results, (a) input of IDRiD dataset image (b) noise removed IDRiD image, (c) contrast enhanced IDRiD image (d) segmented IDRiD image (e) input of RFMiD dataset image (f) noise removed RFMiD image, (g) contrast enhanced RFMiD image (h) segmented RFMiD image (i) input MESSIDOR dataset image (j) noise removed MESSIDOR image, (k) contrast enhanced MESSIDOR image (l) segmented MESSIDOR image (m) input OCT dataset image (n) noise removed OCT image, (o) contrast enhanced OCT image (p) segmented OCT image.

C. Comparative analysis

This section illustrates the comparison of proposed and existing approaches. In this research, we consider three existing works for comparison namely DFTSA [36] CenterNet [37], and OCT images [38]. In addition, we compared our proposed work with respect to four types of datasets such as IDRiD, Fundus-OCT image, RFMiD, and MESSIDOR. The proposed work performance is evaluated by various types of performance metrics such as accuracy, precision, recall, f-measure, computation time, and ROC curve.

a. Impact of Accuracy

This metric is used to evaluate the proposed work accuracy in DR and DME detection. The accuracy is calculated based on the sum of true positive and true negative samples are divided by total number of samples. The calculation of accuracy is defined as follows,

$$A = \frac{P_T + N_T}{P_T + N_T + P_F + N_F} \quad (39)$$

where P_T , N_T represents true positive and true negative respectively. P_F represent false positive and N_F indicates false negatives. Fig 8 depicts the comparison of proposed and existing approaches in terms of accuracy with respect to number of images. The figure shows that the proposed work achieved superior performance compared to existing approaches. The proposed work performs two stages of preprocessing such as noise removal and contrast enhancement which increases high accuracy because of high quality images. Here, we extract only green channel information because of its high contrast which increases the detection accuracy. The existing works does not focus on these two-stage preprocessing, hence it provides low quality images which leads to misclassification that results in less accurate results in DR and DME detection. In addition, we have considered more features for both segmentation and DR, DME detection

which increases the detection accuracy compared to existing approaches. The proposed work achieves 11% higher than DTFSA, and 16.2% higher than CenterNet, and 21.2% higher than OCT images work. Similarly, fig 9 depicts the comparison of accuracy for various datasets. The comparison results demonstrate that the proposed IDRiD dataset achieved superior performance compared to other three datasets.

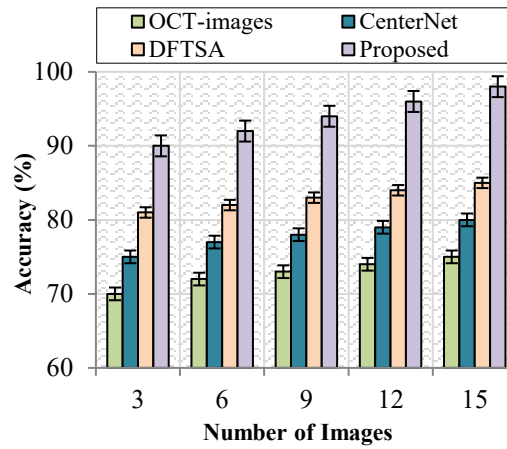


Fig.8 Accuracy vs. Number of Images

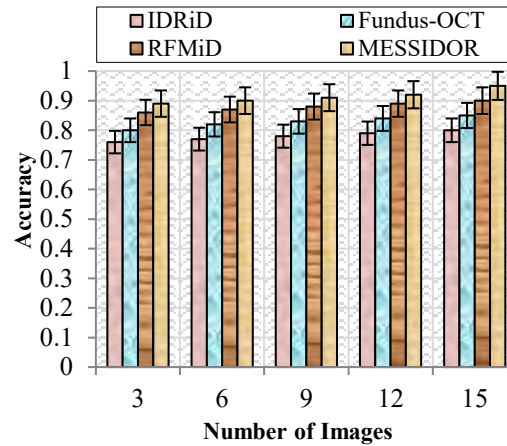


Fig.9 Comparison of Accuracy for Datasets

b. Impact of Precision

This metric helps to evaluate the positive prediction value of the proposed and existing works. The mathematical representation of precision is defined as follows,

$$P = \frac{P_T}{P_T + N_T} \quad (40)$$

The precision is defined as the true positive rate is divided by sum of true positive and true negative. Fig 10 depicts the comparison of precision with respect to number of images. The precision is gradually increases with the increasing number of images. The proposed work achieved high precision when compared to other three existing approaches. In our work, two stage preprocessing is performed to increase the quality of the images. In addition, augmentation is performed to increases the training and prediction efficiency. For doing augmentation, we have multiple view information which increases the true positive rate in DR and DME detection. The existing works does not perform augmentation which reduces the true positive rate due to lack of images in training and testing. Here, we perform both optic disc and lesion segmentation for detecting DR and DME accurately. The existing CenterNet and DFTSA extract the features from the overall image not focusing on particular region or affected region which increases complexity during detection and classification. Additionally, it increases high false positive rate and less true positive rate because of high number of features extracted from overall images. Therefore, the existing work has less precision compared to proposed approach. The proposed work achieves 15% higher than DFTSA, and 20% higher than CenterNet, and 23% higher than OCT images. Similarly, fig 11 represent the precision comparison for various datasets. The numerical result shows that the proposed IDRiD dataset achieves better performance compared to RFMiD, Fundus-OCT images, and MESSIDOR.

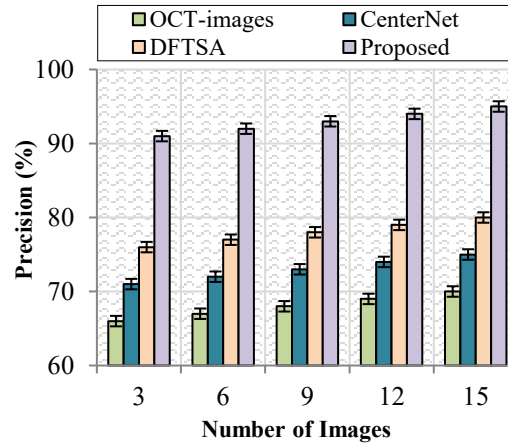


Fig.10 Precision vs. Number of Images

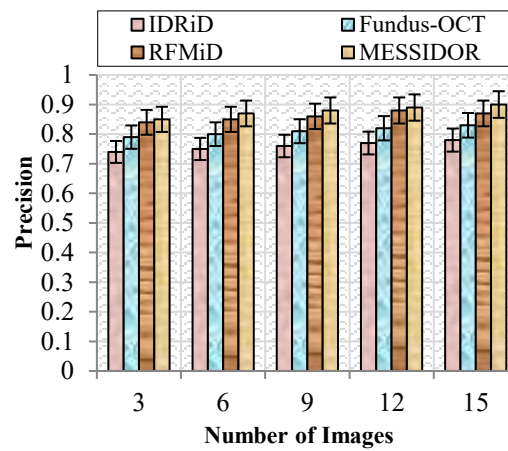


Fig.11 Comparison of Accuracy for Datasets

c. Impact of Recall

This metric is used to detect the correctly recognized true positive rates in DR and DME detection. The definition of recall is defined as the true positive rate is divided by the sum of false negatives and true positive. The mathematical equation of recall is explained as follows,

$$\tilde{R} = \frac{P_T}{P_T + N_F} \quad (41)$$

Fig 12 represents the comparison of recall for both existing (OCT images, CenterNet, DFTSA), and proposed work with respect to number of images. Similarly, fig 13 depicts the comparison of recall for various datasets used in the proposed work. The IDRiD public dataset achieves high recall compared to other three datasets.

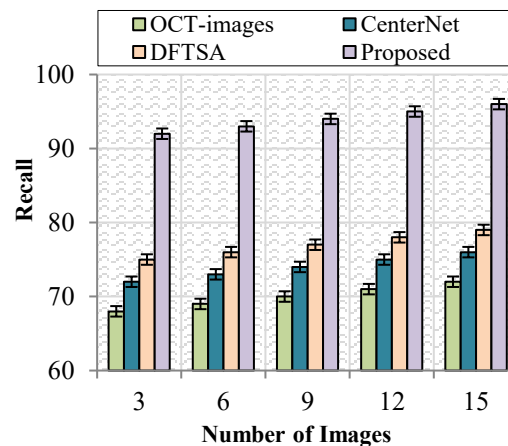


Fig.12 Recall vs. Number of images

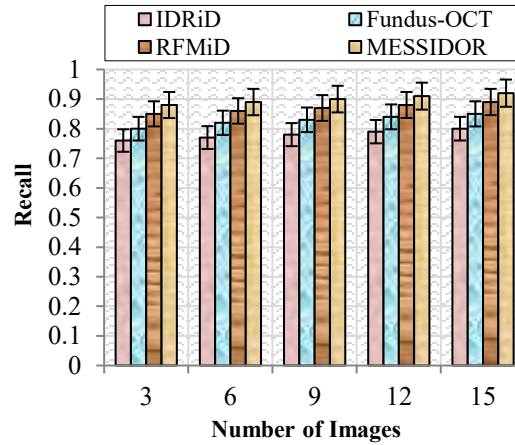


Fig.13 Comparison of Recall for Datasets

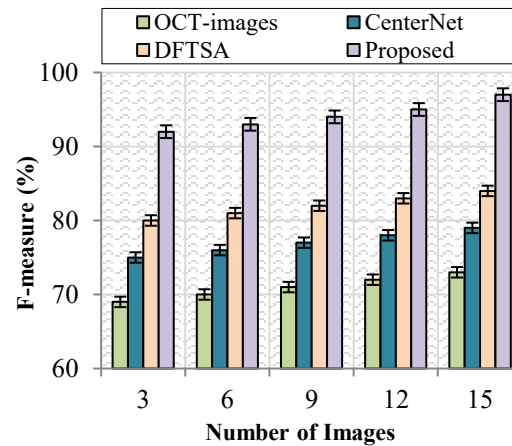
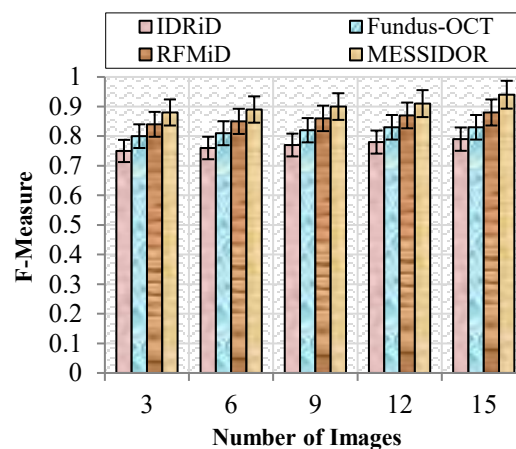
The numerical result shows that the proposed work achieved better recall rate compared to existing approaches. In this research, we perform two stages of preprocessing, data augmentation, feature extraction and classification for increasing true positive rate and reducing false positive rate. Data augmentation increases the training and testing efficiency and detection accuracy due to considering various degrees of images. Segmentation is performed by UNet³⁺ by considering various metrics for both optic disc and lesion segmentation. Existing works only focused on lesion or disc segmentation which leads to misclassification that reduces the true positive rate which also reduces precision and recall. In DR and DME detection we proposed AHS-VGG16 algorithm which consider both low level and high level features for classifying the segmented images, but existing works consider only limited features for DR and DME detection which increases false positive rate and reduces true positive rate. Hence it achieves less recall rate compared to proposed work. The proposed work achieves 17% higher than DFTSA, 20% higher than CenterNet, and 24% higher than OCT images.

d. Impact of F-Measure

The F-measure is calculated based on harmonic mean of the precision and recall. The mathematical equation of the f-measure is formulated as below,

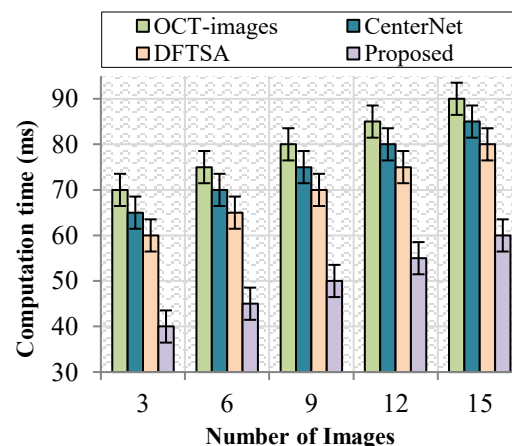
$$\hat{F} = 2 \times \left(\frac{\hat{P} \times \hat{R}}{\hat{P} + \hat{R}} \right) \quad (42)$$

Fig 14 represents the comparison of f-measure for both proposed and existing approaches. Similarly, fig 15 represent the comparison of f-measure for individual datasets. In this research, IDRiD dataset achieves better f-measure compared to other datasets. The numerical result shows that the proposed work achieved superior performance compared to existing approaches. If the work has high precision and recall, then it automatically has high f-measure. The f-measure value is increases gradually with the increasing number of images. The reason of high f-measure is green channel extraction-based contrast enhancement, multi feature based segmentation and classification. In this research, we extract green channel information because of its high contrast. The microneurysm is clearly visible in green channel which increases the detection accuracy, and f-measure. Segmentation also increases the f-measure, because we ensure the segmentation accuracy by validating the similarity using Chi-square which increases segmentation efficiency and f-measure. Here, we perform two segmentation such as optic disc and lesion segmentation for accurate detection of DR and DME. The detection and classification of DR and DME is performed by AHS-VGG-16 which increases f-measure by extracting various features from the segmented images. The proposed work achieves 14.2 % higher compared to DFTSA, and 17.1% greater than CenterNet and 23.2% higher than OCT images.


Fig.14 F-measure vs. Number of images

Fig.15 Comparison of f-measure for Datasets

e. Impact of Computation time

This metric is used to calculate the overall time taken by the work to complete all the processes for detecting DR and DME. The low computation time increases the performance and efficiency of the work. Fig 16 represents the comparison of computation time with respect to number of images. Fig 17 depicts the comparison of computation time for various types of datasets which proved that the IDRiD dataset achieves less computation results compared to other three datasets. From the numerical analysis we proved that the proposed work achieved less computation time compared to existing works.


Fig.16 Computation time vs. Number of images

In this research, we used HBO algorithm for green channel extraction-based contrast enhancement which provides contrast enhancement results in a faster manner due to high convergence. The existing DFTSA model used CLAHE algorithm for contrast enhancement which takes much time for update the clip limit

that increases computation time compared to our proposed HBO algorithm. In addition, DFTSA and OCT images works used CNN algorithm for feature extraction and classification which leads to high computation time due to unwanted generation of convolutional layer during feature extraction and classification. The CenterNet work used centernet algorithm for feature extraction and classification which takes much time for training and testing for large dataset that leads to high computation time. In our work, segmentation is performed before feature extraction and classification which reduces the complexity and computation time during detection, but the existing work extracts the features from the overall imaged rather than extract affected region which leads to high complexity and high computation time compared to our work. The proposed work achieves 20ms less than DFTSA model, and 25ms less than CenterNet model, and 30ms less than OCT images model.

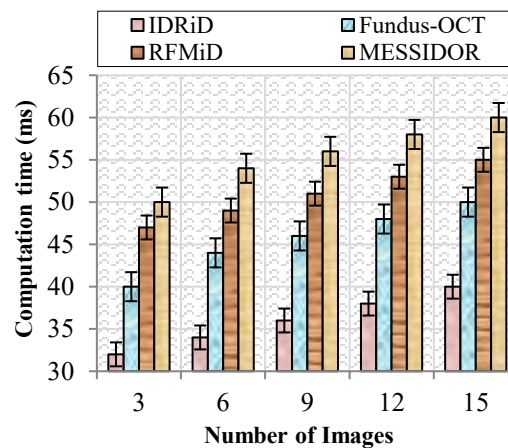


Fig.17 Comparison of Computation time for Datasets

e. Impact of ROC Curve

ROC curve denotes Received Operating Curve that depicts the graphical illustration between true positive rate and false positive rate. In addition, ROC curve shows the accuracy of the works.

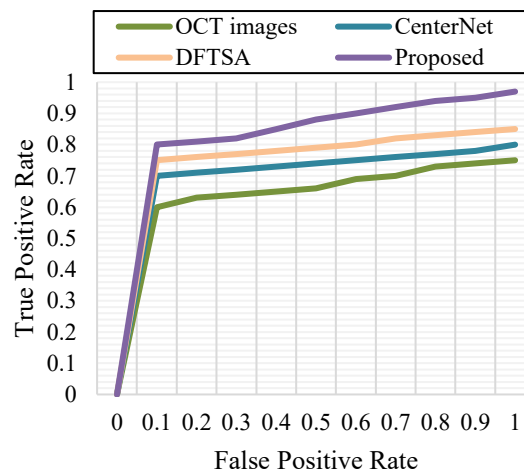


Fig.18 True Positive Rate Vs. False Positive Rate.

The relationship of false positive and true positive rate is shown in fig 18 which proved that the proposed work achieved high performance in ROC compared to existing approaches. The high true positive rate is achieved by performed image quality assessment, bi-level segmentation and multi-feature-based DR and DME detection. The proposed work achieves high true positive rate compared to DFTSA, OCT images, and CenterNet models. If the work has high positive prediction rate, hence it achieves high true positive rate. The proposed work achieved 0.97 true positive rate, and DFTSA achieves 0.85 true positive rate, and CenterNet achieves 0.8 true positive rate, and 0.75 achieves true positive rate.

D. Research Summary

This section summarizes the experimental results of the proposed work. The proposed work implemented by using Matlab simulation tool and proved that the proposed work achieves superior performance in various performance metrics. The performance of this research is shown in fig 8-fig 18. The table illustrates the average numerical values of the performance metrics such as accuracy, precision, recall, f-measure, computation time, and ROC curve for both proposed and existing works. Table illustrates the average numerical values of performance metrics for four various datasets. Table 6 represent the numerical

analysis of proposed and existing approaches, and table 7 represent the numerical analysis of the datasets, both tables are including the average numerical values of the performance metrics.

Table.6 Numerical Analysis of Proposed and Existing approaches

Metrics	IDRiD	Fundus-OCT	RFMiD	MESSIDOR
Accuracy (%)	0.79	0.83	0.88	0.92
Precision (%)	0.76	0.81	0.86	0.88
Recall (%)	0.78	0.83	0.87	0.9
F-Measure (%)	0.77	0.81	0.86	0.90
Computation time (ms)	36	45.6	51	55.6

Table.7 Numerical Analysis of Various Datasets

Metrics	OCT images	CenterNet	DFTSA	Proposed
Accuracy (%)	72.8	77.8	83	94
Precision (%)	68	73	77	93
Recall (%)	70	74	78	94
F-Measure (%)	71	77	82.1	94.2
Computation time (ms)	80	75	70	50
ROC curve	0.61	0.67	0.72	0.80

The highlights of this research are listed as follows,

- For improving the quality of the images, we proposed Image Quality Assessment which includes two processes such as preprocessing, and augmentation which increases detection accuracy and true positive rate.
- For increasing the accuracy of the DR and DME detection we proposed bi-level segmentation using U-NET³⁺ which segments both optic disc and lesions in the retinal images which reduces the complexity of feature extraction and classification.
- For improving the accuracy of detection, we proposed multi-feature-based detection is performed by considering multi-features (i.e low-level and high-level) using Attention-based Hybrid SqueezeNet-VGG16 algorithm which perform feature extraction, fusion and classification for DR and DME detection.

6. CONCLUSION AND FUTURE WORK

In this research, novel deep learning method is proposed for detecting DR and DME using retinal images. Initially, preprocessing is performed by improving the quality of the images, and data augmentation is performed to increase the performance of training and testing which increases the detection accuracy. Here, optic disc and lesion segmentation is performed by considering diameter and color, size, intensity, shape, position respectively by using U-Net³⁺ algorithm which increases segmentation performance because of its novel skip connections. The performance of segmentation is evaluated by Chi-Square which predicts the segmentation error to improve its performance. Finally, DR and DME detection is performed by considering both low level and high-level features which increases detection accuracy. Finally, severity grading is performed for DR by considering lesion, count, location, and exudates thickness which grades the DR into four levels such as mild, moderate, severe, and PDR. The simulation of this research is conducted by Matlab simulation tool which proved that the proposed work achieved superior performance in terms of accuracy, precision, recall, f-measure, computation time, and ROC curve. In future, we planned to detect various retinal image diseases using deep learning approach and the performance is evaluated by various datasets.

REFERENCES

- [1]. S., V., & R, V. (2021). A Survey on Diabetic Retinopathy Disease Detection and Classification using Deep Learning Techniques. 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), 1-4.
- [2]. Rajput, G.S., Reshmi, B.M., & Rajesh, I.S. (2020). Automatic detection and grading of diabetic maculopathy using fundus images. *Procedia Computer Science*, 167, 57-66.
- [3]. Sakthi sree devi, M., Ramkumar, S., Vinuraj kumar, S., & Sasi, G. (2021). Detection of diabetic retinopathy using OCT image. *Materials Today: Proceedings*.
- [4]. Zang, P., Gao, L., Hormel, T., Wang, J., You, Q., Hwang, T.S., & Jia, Y. (2021). DcardNet: Diabetic Retinopathy Classification at Multiple Levels Based on Structural and Angiographic Optical Coherence Tomography. *IEEE Transactions on Biomedical Engineering*, 68, 1859-1870.
- [5]. Maurya, P.K., Gupta, V., Singh, M., Singh, A.K., Kumar, B., & Mohan, A. (2020). Automated detection of diabetic macular edema involving cystoids and serous retinal detachment. *Optics and Laser Technology*, 127, 106157.

- [6]. Mateen, M., Wen, J., Nasrullah, Song, S., & Hayat, S. (2020). Exudate Detection for Diabetic Retinopathy Using Pretrained Convolutional Neural Networks. *Complex.*, 2020, 5801870:1-5801870:11.
- [7]. Kumar, E.S., & Bindu, C.S. (2021). Segmentation of Retinal Lesions in Fundus Images: A Patch Based Approach Using Encoder-Decoder Neural Network. 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 1, 1247-1253.
- [8]. Praveena, S.M., & Lavanya, R. (2019). Superpixel based Segmentation for Multilesion Detection in Diabetic Retinopathy. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 314-319.
- [9]. Maya, K., & Adarsh, K.S. (2019). Detection of Retinal Lesions Based on Deep Learning for Diabetic Retinopathy. 2019 Fifth International Conference on Electrical Energy Systems (ICEES), 1-5.
- [10]. Reddy, S.S., Sethi, N., Rajender, R., & Mahesh, G. (2020). Extensive analysis of machine learning algorithms to early detection of diabetic retinopathy. *Materials Today: Proceedings.*
- [11]. Abdelsalam, M.M., & Zahran, M.A. (2021). A Novel Approach of Diabetic Retinopathy Early Detection Based on Multifractal Geometry Analysis for OCTA Macular Images Using Support Vector Machine. *IEEE Access*, 9, 22844-22858.
- [12]. Deng, J., Tang, P., Zhao, X., Pu, T., Qu, C., & Peng, Z. (2022). Local Structure Awareness-Based Retinal Microaneurysm Detection with Multi-Feature Combination. *Biomedicines*, 10.
- [13]. Ashir, A.M., Ibrahim, S., Abdul-Ghani, M.A., Abdu, A.I., & Anwar, M.S. (2021). Diabetic Retinopathy Detection Using Local Extrema Quantized Haralick Features with Long Short-Term Memory Network. *International Journal of Biomedical Imaging*, 2021.
- [14]. Khan, Z., Khan, F.G., Khan, A., Rehman, Z.U., Shah, S., Qummar, S., Ali, F., & Pack, S. (2021). Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture. *IEEE Access*, 9, 61408-61416.
- [15]. Chen, W., Yang, B., Li, J., & Wang, J. (2020). An Approach to Detecting Diabetic Retinopathy Based on Integrated Shallow Convolutional Neural Networks. *IEEE Access*, 8, 178552-178562.
- [16]. Mobeen-ur-Rehman, Khan, S.H., Abbas, Z., & Danish Rizvi, S. (2019). Classification of Diabetic Retinopathy Images Based on Customised CNN Architecture. 2019 Amity International Conference on Artificial Intelligence (AICAI), 244-248.
- [17]. Zehra, F., Faran, M., Anjum, A., & Islam, S. (2019). DR-Net: CNN Model to Automate Diabetic Retinopathy Stage Diagnosis. 2019 International Conference on Electrical, Electronics and Computer Engineering (UPCON), 1-5.
- [18]. Luo, X., Pu, Z., Xu, Y., Wong, W.K., Su, J., Dou, X., Ye, B., Hu, J., & Mou, L. (2021). MVDRNet: Multi-view diabetic retinopathy detection by combining DCNNs and attention mechanisms. *Pattern Recognit.*, 120, 108104.
- [19]. Tu, Z., Gao, S., Zhou, K., Chen, X., Fu, H., Gu, Z., Cheng, J., Yu, Z., & Liu, J. (2020). SUNet: A Lesion Regularized Model for Simultaneous Diabetic Retinopathy and Diabetic Macular Edema Grading. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 1378-1382.
- [20]. Thulkar, D., Daruwala, R., & Sardar, N.K. (2020). An Integrated System for Detection Exudates and Severity Quantification for Diabetic Macular Edema. *Journal of Medical and Biological Engineering*, 40, 798-820.
- [21]. Thulkar, D., Daruwala, R., & Sardar, N.K. (2020). An Integrated System for Detection Exudates and Severity Quantification for Diabetic Macular Edema. *Journal of Medical and Biological Engineering*, 40, 798-820.
- [22]. Wang, Z., Zhang, W., Sun, Y., Yao, M., & Yan, B. (2020). Detection of Diabetic Macular Edema in Optical Coherence Tomography Image Using an Improved Level Set Algorithm. *BioMed Research International*, 2020.
- [23]. Gao, J., Qummar, S., Zhang, J., Yao, R., & Khan, F.G. (2020). Ensemble Framework of Deep CNNs for Diabetic Retinopathy Detection. *Computational Intelligence and Neuroscience*, 2020.
- [24]. Pao, S., Lin, H., Chien, K., Tai, M., Chen, J., & Lin, G. (2020). Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network. *Journal of Ophthalmology*, 2020.
- [25]. Mateen, M., Malik, T.S., Hayat, S., Hameed, M., Sun, S., Wen, J. (2022). – Deep Learning Approach for Automatic Microaneurysms Detection. *Sensors* 22, 542.
- [26]. Qummar, S., Khan, F.G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z.U., Ahmed Khan, I., & Jadoon, W. (2019). A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. *IEEE Access*, 7, 150530-150539.
- [27]. Tavakoli, M., Mehdizadeh, A., Aghayan, A., Shahri, R.P., Ellis, T.J., & Dehmeshki, J. (2021). Automated Microaneurysms Detection in Retinal Images Using Radon Transform and Supervised Learning: Application to Mass Screening of Diabetic Retinopathy. *IEEE Access*, 9, 67302-67314.
- [28]. Saeed, F., Hussain, M., & Aboalsamh, H.A. (2021). Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network. *IEEE Access*, 9, 41344-41359.
- [29]. Bilal, A., Sun, G., Li, Y., Mazhar, S., & Khan, A.Q. (2021). Diabetic Retinopathy Detection and Classification Using Mixed Models for a Disease Grading Database. *IEEE Access*, 9, 23544-23553.
- [30]. Vidal, P.F., Moura, J.D., Díaz, M., Novo, J., & Ortega, M. (2020). Diabetic Macular Edema Characterization and Visualization Using Optical Coherence Tomography Images. *Applied Sciences*, 10, 7718.

- [31]. Wan, C., Chen, Y., Li, H., Zheng, B., Chen, N., Yang, W., Wang, C., & Li, Y. (2021). EAD-Net: A Novel Lesion Segmentation Method in Diabetic Retinopathy Using Neural Networks. *Disease Markers*, 2021.
- [32]. Nahiduzzaman, M., Islam, M.R., Islam, S.M., Goni, M.O., Anower, M.S., & Kwak, K.S. (2021). Hybrid CNN-SVD Based Prominent Feature Extraction and Selection for Grading Diabetic Retinopathy Using Extreme Learning Machine Algorithm. *IEEE Access*, 9, 152261-152274.
- [33]. Yasin, S., Iqbal, N., Ali, T., Draz, U., Alqahtani, A., Irfan, M., Rehman, A., Głowacz, A., Alqhtani, S.M., Proniewska, K., Brumercik, F., & Wzorek, L. (2021). Severity Grading and Early Retinopathy Lesion Detection through Hybrid Inception-ResNet Architecture. *Sensors (Basel, Switzerland)*, 21.
- [34]. Singh, R.K., & Gorantla, R. (2020). DMENet: Diabetic Macular Edema diagnosis using Hierarchical Ensemble of CNNs. *PLoS ONE*, 15.
- [35]. Yaqoob, M.K., Ali, S.F., Bilal, M., Hanif, S.M., & Al-Saggaf, U.M. (2021). ResNet Based Deep Features and Random Forest Classifier for Diabetic Retinopathy Detection †. *Sensors (Basel, Switzerland)*, 21.
- [36]. Atteia, G., Samee, N.M., & Hassan, H.Z. (2021). DFTSA-Net: Deep Feature Transfer-Based Stacked Autoencoder Network for DME Diagnosis. *Entropy*, 23.
- [37]. Nazir, T., Nawaz, M., Rashid, J., Mahum, R., Masood, M.F., Mehmood, A., Ali, F., Kim, J., Kwon, H., & Hussain, A. (2021). Detection of Diabetic Eye Disease from Retinal Images Using a Deep Learning based CenterNet Model. *Sensors (Basel, Switzerland)*, 21.
- [38]. Ghazal, M., Ali, S., Mahmoud, A.H., Shalaby, A.M., & El-Baz, A.S. (2020). Accurate Detection of Non-Proliferative Diabetic Retinopathy in Optical Coherence Tomography Images Using Convolutional Neural Networks. *IEEE Access*, 8, 34387-34397.
- [39]. Nazir, T., Irtaza, A., Javed, A., Malik, H., Hussain, D., & Naqvi, R.A. (2020). Retinal Image Analysis for Diabetes-Based Eye Disease Detection Using Deep Learning. *Applied Sciences*, 10, 6185.
- [40]. Patil, S.B., & Patil, B. (2020). Automated macula proximity diagnosis for early finding of diabetic macular edema. *Research on Biomedical Engineering*, 1-17.