A Novel Deep Learning Framework for Diagnosing Retinal Diseases using OCT images

[1]Sathish Kamalakannan, [2]Dr.B.Kirubagari, [3]Dr.J.Jegan
[1]Research Scholar, [2]Professor, [3]Assistant Professor
[12]Dept of Computer Science and Engineering, [3]Faculty of Engineering &Technology
[12]Faculty of Engineering &Technology, [3]SRM Institute of Science and Technology Tiruchirapalli

Abstract—**Ophthalmologists employ optical coherence tomography (OCT), a non-invasive imaging technique that produces cross-sectional images of the retinal layers, to diagnose retinal abnormalities. OCT is a crucial tool for diagnosing and assessing retinal diseases. However, analyzing the many images that OCT generates for each patient often takes a substantial amount of time for ophthalmologists. In this work, OCT images are classified into four groups using deep learning models: normal, drusen, diabetic macular edema (DME), and choroidal neovascularization (CNV). Two distinct models are proposed: one employs segmentation based on curvature-based regions of interest (ROI), while the other uses the Folded RESNET101 method for classification. The recommended method's accuracy, sensitivity, and specificity are all 0.99, producing outstanding outcomes.**

Index Terms—**Deep Learning, Convolution Neural Network, OCT**

I. Introduction

Using a non-invasive imaging technique called OCT, cross-sectional images of the retina layer are produced. It takes use of near-infrared light pulses, which may penetrate the retinal layer as deep as several hundred microns. Therefore, it is essential for ophthalmologists to detect and quantify retinal abnormalities and diseases and to recommend treatment for conditions such as glaucoma, diabetic macular edoema (DME), age-related macular degeneration (AMD), choroidal neo-vascularization (CNV), and diabetic retinopathy (DR) [1, 2]. There are two subtypes of AMD: wet AMD and dry AMD. The majority of dry AMD patients have Drusen, while wet AMD patients typically have Choroidal Neovascularisation (CNV) and associated retinal symptoms. [3] [4] [5]. CNV is the formation of abnormal blood vessels in the retina's choroid layer [6]. DME is a fluid accumulation caused by blood vessel leakage in the macula region. Thirty percent of DR patients experience DME [7, 8]. Since each patient obtains several images from the process, ophthalmologists must invest a significant amount of time in OCT analysis. Deep learning methods have recently been adapted for OCT image classification. DenseNet201 is used to categorise OCT into four types [9] [10]. Using Inception V3 [11], OCT images are divided into four categories: Normal, Drusen, DME, and CNV [12][18]. Ensemble based learning on ResNet152 [13] is used to classify OCT images into the four classes [14]. Using modified ResNet50 and ensemble learning, the photos were categorised into four classes [13] [15][17]. VGG16 and image normalisation are used to train OCT pictures for the four-class categorisation [16][19].

The aforementioned investigations employed two different deep learning (DL) models. One uses multiple binary classifiers, while the other uses multi-class classifiers. Since only one classifier needs to be trained, the DL model employing a multi-class classifier (DLM) seems more practical for handling multi-class classification issues than the DL model using binary classifiers (DLB).

For each class, the DLB must train several classifiers. There is a benefit to using the DLB when we update or improve the model to classify one or more categories in the same data after the model has been constructed. For new classes, the DLB just needs to train more binary classifiers. Nevertheless, the DLM must retrain the model.The DLB has to train multiple classifiers for every class. When we update or enhance the model to classify one or more categories in the same data after the model has been built, there is a benefit to employing the DLB. The DLB only needs to train additional binary classifiers for new classes. However, the model needs to be retrained by the DLM. The DLB has greater extensibility than the DLM. Multiple binary classifiers perform better than a single multi-class classifier in the picture segmentation scenario [20]. This study proposes two DL models to classify patients' OCT pictures into four groups: Normal, DME, CNV, and Drusen. Four binary CNN classifiers are used in the design of the suggested models' architectures.

A preprocessing technique is used to trim the retinal layers and remove noise from the pictures. The following is an explanation of the remainder of this work. The data and our categorisation methods for OCT pictures are covered in detail in Section II. The results of the experiments are described in Section III, and the conclusion is presented in Section IV.

II. PROPOSED METHOD

Pre-processing, image segmentation, feature extraction, and classification are the steps of the recommended method. The stages involved in the suggested method are detailed in figure 1 below.

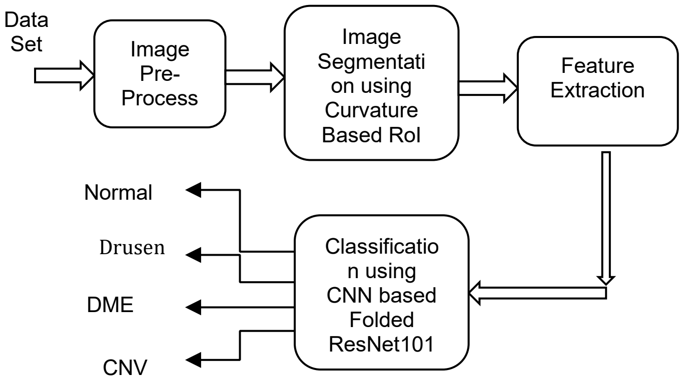


Figure 1 : Architecture Diagram for Proposed method

A.IMAGE PRE-PROCESSING:

Since Gaussian Filter and Pre-ROI techniques are more effective at minimising noise distortion and have a more realistic appearance among the visualization-based family, they are used throughout this study to enhance contrast and picture features by highlighting anomalies.

i) Apply Gaussian filter to reduce the noise

A Gaussian filter is a low pass filter used to reduce noise (high frequency components) and blur hard areas of an image. To get the desired outcome, the filter is constructed as a straightforward array of odd-sized symmetrical kernels that are processed through each pixel in the area of interest. The kernel is not sensitive to sudden colour changes (edges) since its inner pixels contribute far more to the final value than any of its outside ones. A Gaussian Filter can be thought of as a rough representation of the Gaussian Function.

$$g(x,y) = 1/2\pi\sigma^2[e^{-(x2+y2)/(2\sigma2)}]$$

x-row and y-column

i) Pre- ROI(Region of Interest)

Pre-ROI is used to remove top(1 to 60) and bottom row (450 to end).

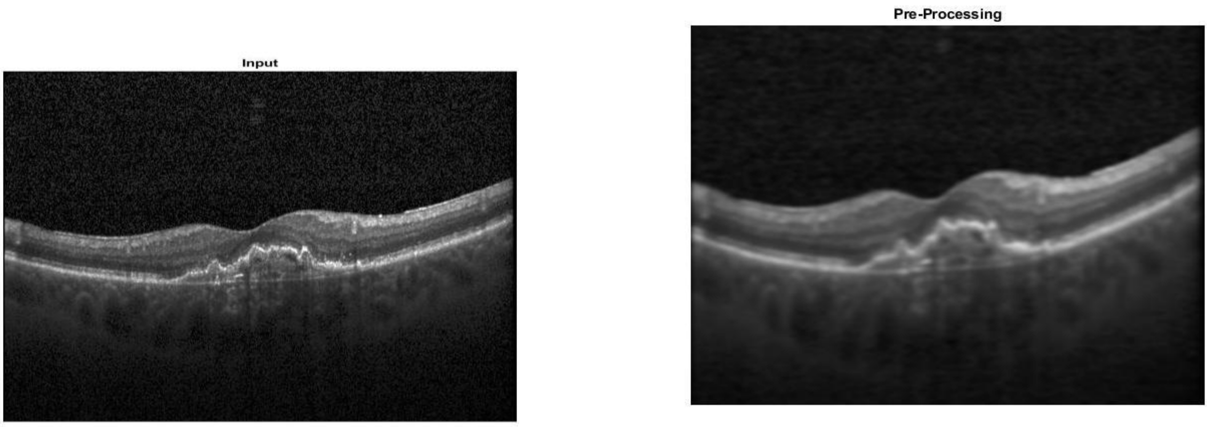


Figure 2 (a) : Input Image (b) : Preprocessed Image

B. PROPOSED SEGMENTATION ALGORITHM – CURVATURE BASED ROI

The proposed method uses curvature based RoI segmentation algorithm is used to extract the features with the following steps.

i) Region_Idendify
Step 1: Make a mask using the image's input size.
Step 2: locate the external and internal points
Step 3: identify the points of curvature
Step 4: Give the points a gradient
Step 5: Verify that the gradient point (0.45) is greater than 0.09
Step 6: complete the image by applying steps 1 through 5.

ii) adjacency matrix:
Step 1: use internal and exterior points to determine the adjacent matrix index.

iii) Get shortest path:
Step 1: To compare the value of the adjascent matrix with its curvature
Step 2: Determine the border values for every column and row.

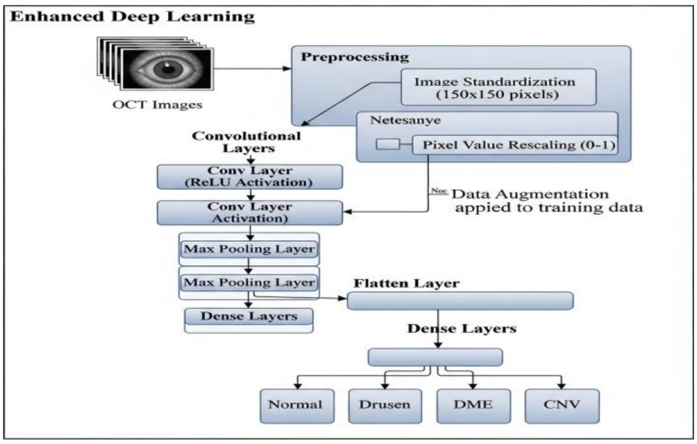


Figure 3 (1-d) displays the input image, initialisation, and image following 1500 iterations, along with a region that has been discovered.

C. THE PROPOSED CLASSIFICATION ALGORITHM

The two deep CNNs used in this study were the Faster CNN and the ResNet101 for classification. The model was constructed and trained using the Python programming language and the Keras package on the Tensor Flow backend. A ResNet101CNN architecture that had previously been trained on a Kaggle database was then retrained on a dataset using transfer learning, which enables an algorithm to leverage accumulated knowledge from previous datasets. The ResNet101's typical architecture is depicted in Figure 5(a).

Algorithm 1: OCT-based Lesion Localization	
Initialize: OCT Dataset DDD Output: Disease label, lesion mask	
Compute	
While (epoch=1) do	

Algorithm 1: OCT-based Lesion Localization			
Initialize: OCT Dataset DDD Output: Disease label, lesion mask			
Compute			
While (epoch=1) do			
	For (each epoch) do		
	Update		
		For (each image $X \in D$) do	
		$Xp \leftarrow Preprocess(X)$	
			$S(X) \leftarrow SegmentationNetwork(Xp)$
			$\{Fi\} \leftarrow Extract\ multi-scale\ features$ $Ffusion \leftarrow AttentionFusion(Fi)$ $FFinal \leftarrow$ Concatenate($FFusion$, S(X)) $Ypred \leftarrow$ Classifier($FFinal$)
	End		End Return disease labels and lesion maps

End	

Algorithm 2 describes the OCT image preprocessing stage. Since raw OCT images often contain noise, intensity variations,and acquisition artifacts, preprocessing is required to enhance image quality and improve feature consistency. In this stage, denoising filters are applied to remove acquisition noise, followed by contrast enhancement using CLAHE to improve retinal structure visibility. The images are then normalized to maintain consistent intensity distribution, and data augmentation techniques are applied to improve model generalization during training.

Algorithm 2: OCT-Image Pre-processing and Enhancement	
Initialize: Raw OCT image X Output: Preprocessed image Xp	
Compute	
While (epoch=1) do Load OCT image X	
For (each epoch) do	
Apply denoising filter to remove acquisition noise	
	For (each image $X \in D$) do
	Perform contrast enhancement using CLAHE Normalize image intensity values Apply data augmentation (rotation, flipping, scaling)
End	End Return enhanced image Xp for segmentation stage.
End	

The CNN recognised and categorised local features in images using its many convolutional layers. It had completely connected levels to aggregate these features and produce a final probability value for the class, as well as pooling layers (average pool and max pool) that combined semantically related features into a single feature.

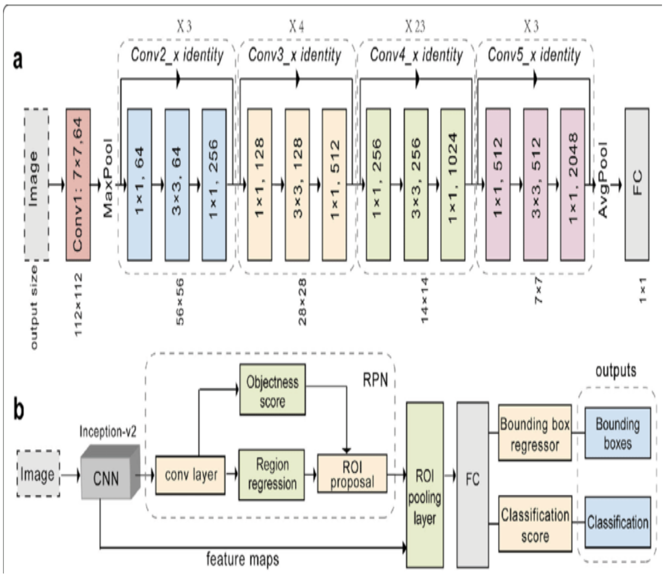


Figure 5: (a) Architecture of Resnet 101 (b) Flow chart of Multiscale image Analysis Model

Throughout the deep network, shortcut links are added every three convolutional layers. These shortcut connections carry out identity mapping without adding more parameters or increasing the computational complexity, which facilitates network optimisation during training. ResNet makes it possible for deeper networks to perform image classification tasks with more accuracy than shallower networks. The model was retrained for the present task after the link weights from the previously trained model were transferred to the new model throughout the training phase. The outputs of the region proposal network, which was used to identify items in the image, were evaluated for potential containers using the outline box regression model and object classifiers. Finally, area of interest pooling and convolution layers were assessed. Lastly, for every target object, the model creates the bounding box and associated category label. After processing a large dataset, a model that combines these two CNN networks will ultimately generate categorisation.

III. EVALUATION METRICS AND RESULT ANALYSIS

In this experiment, optical coherence tomography (OCT) pictures from a Kaggle dataset are used. It is open to the general public. The dataset consists of 4000 test pictures and around 16,000 training photos. There are four courses in it. There are 4000 images in the CNV class, 4000 in the DME class, 4000 in the Drusen class, and 4000 in the Normal class. There are 200 photos in each class in the test dataset.

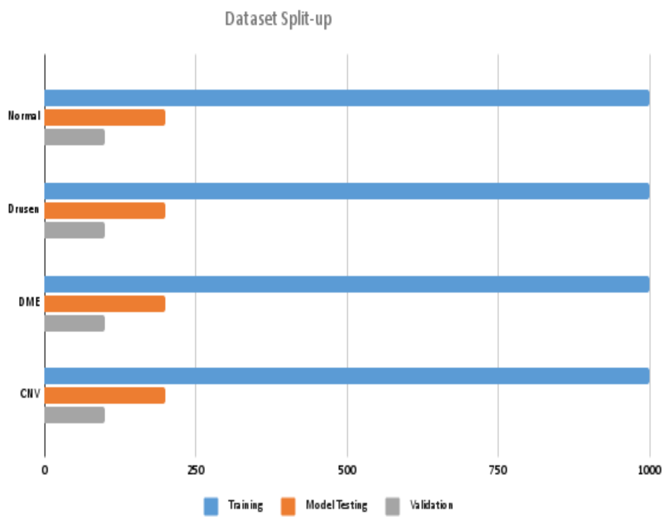
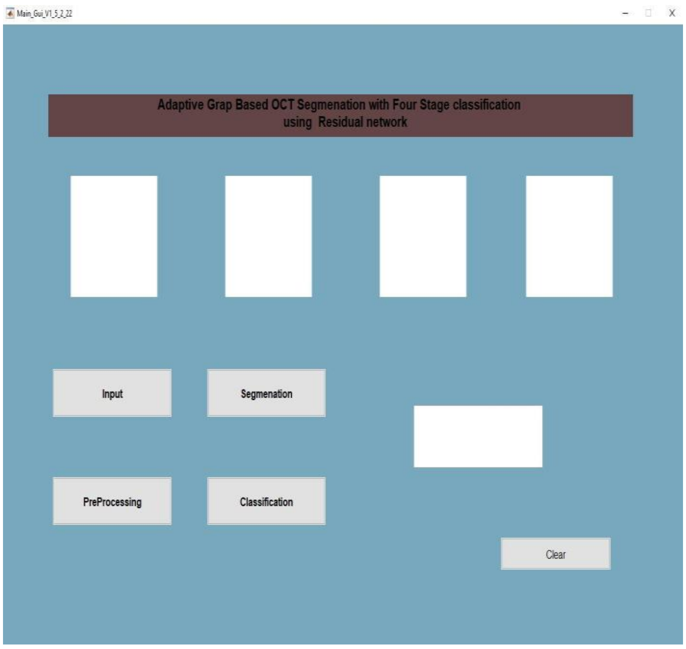


Figure 6: Dataset usage for Multiscale image Analysis method

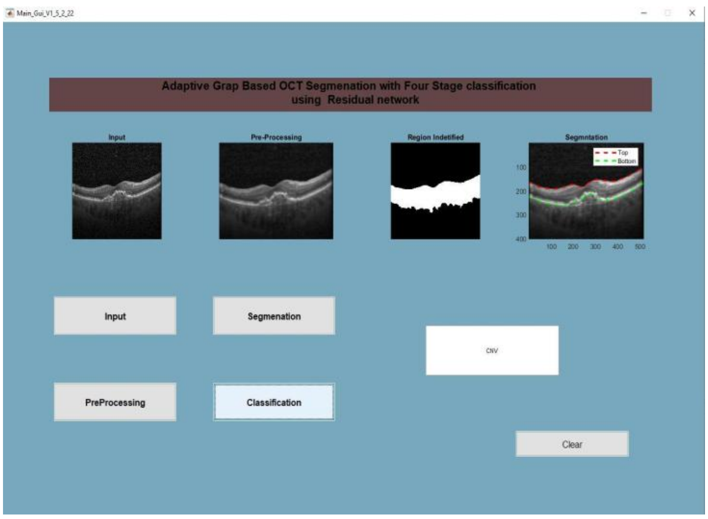
Table 6 and Figure 6 shows the images used in proposed method clearly and the images used for training, testing and validation

Table 1: Dataset Usage for Multiscale image Analysis Method

Metrics used for evaluation include sensitivity, accuracy, specificity, F1 Score, and precision, which are listed in equations 1 through 3. These measurements are measured using four distinct parameters: false-negative (FN), false positive, true positive (TP), and true negative (TN) (FP). The percentage of documents from all the data that were correctly classified is known as accuracy. The percentage of performance that is relevant is called precision. Conversely, recall refers to the proportion of accurately categorised results that the algorithm generated for every functional outcome. The detection rate (DR), sometimes referred to as the true positive rate, is the proportion of anomalous records that are accurately classified as abnormalities to all anomaly data. (TPR) [19, 20].



a. Initial GUI Window



Type	Training	Model Testing	Validation
CNV	4000	200	100
DME	4000	200	100
Drusen	4000	200	100
Normal	4000	200	100

b. Final GUI with segmentation and Classification

Figure 7: (a) Initial GUI window (b) Final GUI window with Segmentation and Classification

The output screenshots of the implementation process are displayed in Figure 7(a) and (b).

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

$$precision = \frac{TP}{FP + TP} \quad (3)$$

$$SP = \frac{TN}{TN + FP} \quad (4)$$

$$F1 - Score = \frac{2 * \Pr ecision * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call}$$

The above said quality measures are taken into consideration for measuring and comparing the performance of the proposed work with the existing model.

Table 2: True Positive, True Negative, False Positive and False Negative Count for Existing and Proposed Method

	Existing Method 1 (OCT Net)	**Existing Method 2 (VGG16)**	**Proposed Method (Folded Resnet101)**
TP	92	94	99
FP	10	8	6
FN	8	7	5
TN	7	6	5

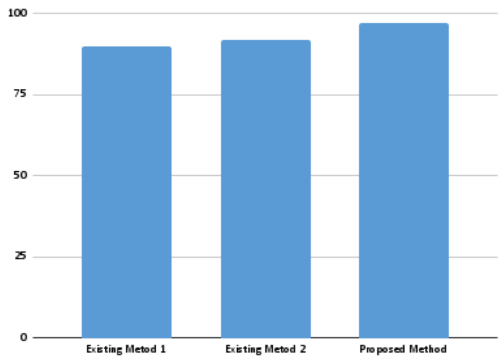


Figure 8: True Positive of Existing and Proposed method.

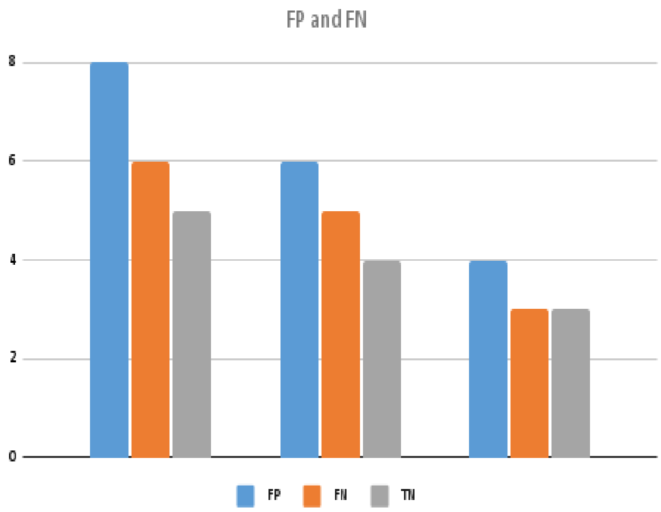


Figure 9: FP,FN, and TN of Existing and Proposed method

The TP, TN, FP, and FN counts of the current and suggested methods are displayed in Table 2, Figures 8, and 9. Compared to other methods now in use, the suggested approach produced a higher result for ZTrue positive. FP, FN, and TN are lower in the suggested approach. This demonstrates that the suggested approach performs better than current approaches.

Table 3: Performance Analysis with respect to existing method

Quality Measures	Existing Method 1(OCT Net)	Existing Method 2(VGG16)	Proposed Method (Folded Resnet101)
Accuracy	88.2	90.7	94.5
Precision	92.8	94.9	97.0
Recall/Sensitivity	94.8	95.8	98.0
Specificity	39.5	41.0	43.9
F1 Score	93.8	95.4	97.5

The performance analysis of current and suggested approaches using a variety of quality metrics, including accuracy, precision, sensitivity, specificity, and F1-score, is displayed in Table 3 and Figure 9. The suggested method has the best F1-Score, accuracy, sensitivity, precision, and specificity. For the Normal class, the binary classifier's accuracy is 93.5%.

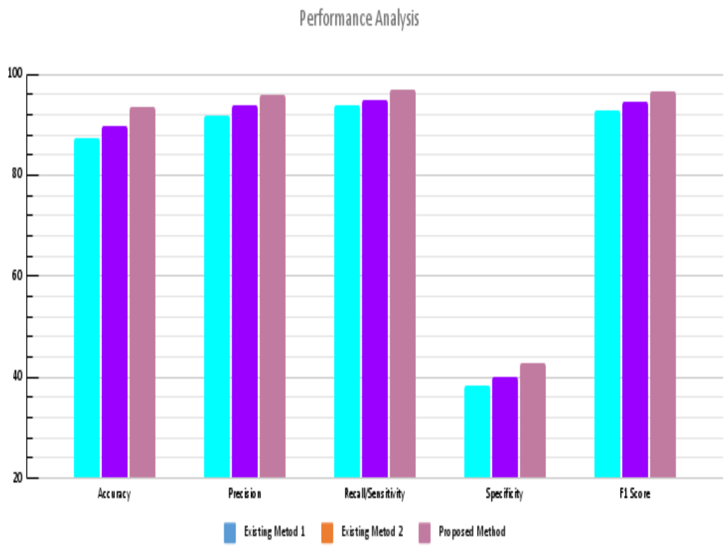


Figure 9: Performance Analysis of Existing and Proposed Method

The class's accuracy rate is 96%. The F1 Score, Specificity, and Recall are 96.5%, 7.9%, and 97%, respectively. These findings demonstrate that, when compared to other current technologies taken into account, the suggested approach performs better in every quality metric. S These results show that they are capable of serving as ophthalmologists' principal method.

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

IV. CONCLUSION

All things considered, our DL-based method showed promise for very accurate and exact automatic recognition of OCT pictures with four-level categorisation. The technology may be used to support clinical judgements because its performance was comparable to or superior to that of retinal specialists. This is the initial phase of the clinical translation of the approach. As this research explains, under the current situation, a system to automatically categorise retinal illness is crucial. The suggested approach is completely capable of detecting retinal disorders in real time. Results confirm that the suggested model can identify general disease kinds in the Kaggle database with 93.5% of extremely high accuracy in real time and offer potential remedies for retinal disorders. This approach can be expanded in the future with a high-performance deep learning model to help ophthalmology specialists categorise the disorders more precisely.

References

[1] Americal Academy of Ophthalmology, https://www.aao.org/eyehealth/treatments/what-is-optical-coherence-tomography

[2] S. Aumann, S. Donner, J. Fischer, and F. Müller, High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics, Springer, August, 2019

[3] Age-Related Macular Degeneration, https://www.nei.nih.gov/learnabout-eye-health/eye-conditions-and-diseases/age-related-maculardegeneration.

[4] Age-related macular degeneration, NICE Guideline, No. 82, National Institute for Health and Care Excellence (UK), January, 2018.

[5] American Macular Degeneration Foundation, https://www.macular.org/dry-vs-wet-macular-degeneration.

[6] A.Umamageswari, N.Bharathiraja, D.Shiny Irene "A Novel Fuzzy C-Means based Chameleon Swarm Algorithm for Segmentation and Progressive Neural Architecture Search for Plant Disease Classification", (2021), ICT ECPRESS, Elsevier, https://doi.org/10.1016/j.icte.2021.08.019M R Kesen and S W Cousins, Choroidal Neovascularization, Encyclopedia of the Eye, Elsevier Ltd., pp. 257-265, 2010.

[7] https://aerpio.com/therapeutic-areas/wet-amddme/

[8] S.R. Cohen and T.W. Gardner, "Diabetic Retinopathy and Diabetic Macular Edema", Dev Ophthalmology, Vol 55, pp. 137-146, 2016

[9] G. Huang, Z Liu, L.V.D. Maaten, K.Q. Weinberger, "Densely Connected Convolutional Networks", CVPR 2017.

[10] K.T. Islam, S. Wijewickrema, and S. O'Leary, "Identifying Diabetic Retinopathy from OCT Images using Deep Transfer Learning with Artificial Neural Networks", 32nd International Symposium on ComputerBased Medical Systems, pp. 281-286, June 2019.

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", IWWW Conference on Cimputer Vision and Pattern Recognition, pp. 2818-2826, 2016.

[12] D.S. Kermany, M. Goldbaum, W. Cai, et al., ''Identifying medical diagnoses and treatable diseases by image-based deep learning,'' Cell, vol. 172, no. 5, pp. 1122–1131, 2018.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, Las Vegas, USA, 2016.

[14] J. Kim and L. Tran, "Ensemble Learning based on Convolutional Neural Networks for the Classification of Retinal Diseases from Optical Coherence Tomography Images," IEEE 33th International Symposium on Computer-Based Medical Systems, pp 535-540, Rochester, USA, July 2020.

[15] F. Li, H. Chen, Z. Liu, X. Zhang, M. Jiang, Z. Wu, and K. Zhou, "Deep Learning-based Automated Detection of Retinal Diseases Using Ophtical Coherence Tomography Images", Biomedical Optics Express, Vol 10, No 12, pp. 6204-6226, December 2019

[16] K, Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", International Conference on Learning Representations, 2015

[17] F. Li, H. Chen, Z Liu, X. Zhang, and Z. Wu, "Fully Automated Detection of Retinal Disorders by Image Based Deep Learning", Gradfe's Archive for Clinical and Experimental Ophthalmology, 257, pp. 495-505, Jaunary 2019

[18] W. Lu, Y. Tong, Y. Yu, Y. Xing, C. Chen, and Y. Shen, "Depp Learning Based Automated Classification of Multi-Categorical Abnormalities From Optical Coherence Tomography Images", Translational Vision Science and Technology (TVST), vol. 7, no. 6, pp.1-10, 2018

[19] A.Umamageswari, J.Shiny Duela, Diolin Sara, K.Raja (2022), “An Enhanced Identification and Classification Algorithm for Plant Leaf Diseases Based on Deep Learning”, Traitement Du Signal, Vol: 39, Issue : 3, pp:1013-1018, https://doi.org/10.18280/ts.390328, (SCIE / SCOPUS Indexed)

[20] A.Umamageswari, S.Deepa, L.Sherin beevi (2022), “A Novel Approach for classification of diabetics from retinal image using deep learning technique”, International Journal Health Sciences, 6(S1), 2729-2736. Http://doi.org/10.53730/ijhs.v6nS1.5196