RETINAL OCT IMAGE SEGMENTATION AND CLASSIFICATION USING DEEP LEARNING

[1]Abinaya R, [2]Sreenidhi S, [3]M Priyadharshini

[1,2,3] Student

[1]Department of Biomedical Engineering,

[1]SRM Institute of Science and Technology, Ramapuram, Chennai, India

[1]rvabikumar@gmail.com, [2] sridharsreenidhi05@gmail.com, [3] priyadharshini.mohan192005@gmail.com

***Abstract*—Automated detection and classification of retinal abnormalities in Optical Coherence Tomography (OCT) images are critical for early diagnosis and effective treatment planning in ophthalmology. In this work, a hybrid methodology is introduced that combines traditional image processing techniques with deep learning for improved segmentation and classification accuracy. The preprocessing pipeline begins with grayscale conversion and contrast enhancement using Adaptive Histogram Equalization (CLAHE). Segmentation is performed using a combination of Otsu's thresholding, morphological operations, and K-Means clustering, followed by refinement using the Active Contour (Snake) algorithm. These steps isolate key regions of interest in the retinal layers, focusing on abnormalities such as fluid-filled areas or distorted tissue structures. From the segmented images, quantitative features including area, eccentricity, perimeter, and circularity are extracted to understand the shape characteristics of the abnormal regions. These segmented images are then used to train a deep learning model based on the GoogleNet architecture, enabling robust classification of retinal conditions. During testing, raw OCT scan images undergo the same segmentation and feature extraction pipeline before classification, ensuring consistency between training and testing datasets. The integrated approach leverages the precision of classical segmentation techniques and the powerful feature learning capabilities of deep convolutional networks. This method enhances the reliability of OCT-based diagnostic systems and can be potentially extended to other forms of medical imaging. The system provides a cost-effective and scalable solution for aiding ophthalmologists in the early detection of retinal diseases such as age-related macular degeneration and diabetic retinopathy.**

***Index Terms*—OCT, segmentation, CLAHE, deep learning, retinal diseases.**

I. INTRODUCTION

Optical Coherence Tomography (OCT) is a non-invasive imaging technique widely used in ophthalmology for capturing high-resolution cross-sectional images of the retina. It plays a vital role in the early detection and monitoring of retinal diseases such as Age-related Macular Degeneration (AMD), Diabetic Retinopathy (DR), and Central Serous Retinopathy. Manual interpretation of OCT images, however, is time-consuming and requires expert knowledge, making it challenging in clinical settings with a high patient load.

To address this, automated methods are increasingly being explored to assist clinicians by providing fast, consistent, and objective analysis of OCT scans. This research introduces a hybrid image analysis framework that combines classical image processing techniques and deep learning for efficient retinal image segmentation and classification. Initially, the OCT image undergoes grayscale conversion and enhancement using Contrast-Limited Adaptive Histogram Equalization (CLAHE) to improve visibility of subtle features. Segmentation is achieved using a combination of Otsu's thresholding, morphological operations, K-Means clustering, and active contour models to isolate abnormal regions accurately.

Quantitative shape analysis is performed on the segmented regions to extract critical features such as area, perimeter, circularity, and eccentricity. These segmented images are then used to train a deep learning model using the pre-trained GoogLeNet architecture. The model is capable of classifying new OCT images based on the learned patterns, even when tested with unsegmented scan images. This hybrid approach improves interpretability, enhances segmentation precision, and leverages the power of transfer learning, making it a promising tool for real-time, AI-assisted ophthalmic diagnosis.

II. LITERATURE REVIEW

Optical Coherence Tomography (OCT) imaging plays a pivotal role in diagnosing retinal conditions such as Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), and Age-related Macular Degeneration (AMD). Traditional methods of OCT analysis have heavily relied on classical image processing techniques including thresholding, edge detection, and morphological operations to delineate anatomical structures and detect pathological regions. For instance, Otsu's thresholding method has been widely used for automatic image binarization based on pixel intensity distributions, and morphological filters have proven effective in removing noise and refining segmentation outputs. Clustering algorithms like K Means have also been applied for partitioning OCT images into meaningful regions based on pixel similarities.

Recent advances in artificial intelligence have introduced deep learning frameworks, particularly Convolutional Neural Networks (CNNs), into OCT image analysis. Pre-trained networks such as GoogLeNet, ResNet, and VGG16 have demonstrated remarkable performance in classifying retinal pathologies with high precision. GoogLeNet, in particular, offers a balanced trade-off between computational efficiency and accuracy by utilizing inception modules to capture multi-scale features. Several studies have leveraged transfer learning to adapt these models to medical imaging, where annotated data is often scarce. Hybrid approaches that integrate traditional segmentation techniques with deep learning classification are increasingly gaining attention. These methods aim to combine the localization power of conventional algorithms with the high-level feature representation of neural networks. For example, segmenting regions of interest using K-Means clustering or active contours, followed by classification with GoogLeNet, has been shown to improve diagnostic performance. Such integrated methodologies are paving the way for more robust and interpretable automated retinal analysis systems in clinical settings.

Numerous studies have highlighted the importance of precise segmentation in improving the accuracy of OCT image classification. Researchers have explored the use of active contour models and region-based segmentation as a preprocessing step to isolate pathological regions. These segmented areas, when fed into deep learning networks like GoogLeNet, have been shown to enhance classification outcomes by focusing the model's attention on disease-relevant features. For instance, a study by Kermany et al. demonstrated that combining image pre-processing with deep CNN architectures significantly improved the identification of retinal diseases. Their work laid the foundation for subsequent research combining medical image segmentation with transfer learning models. Other approaches have incorporated unsupervised learning techniques such as clustering to pre-label or extract features from OCT scans before classification. K Means clustering, in particular, has been frequently used for distinguishing retinal layers or identifying fluid-filled regions, which are indicative of retinal pathologies.

Studies have shown that hybridizing these clustering results with supervised learning enhances the model's ability to differentiate between subtle texture changes in normal and abnormal retinal structures. Recent literature also emphasizes the need for automated systems that can generalize across different patient populations and imaging devices. Transfer learning with GoogLeNet addresses this challenge by leveraging features learned from large scale image databases and adapting them to OCT datasets. This not only reduces the need for extensive training data but also improves the robustness of the model. Collectively, these findings support the integration of classical segmentation with deep learning classification as an effective strategy for advancing retinal disease diagnosis.

III. METHODOLOGY

The methodology involves a hybrid approach combining classical image processing techniques and deep learning-based classification for the analysis of Optical Coherence Tomography (OCT) images. The process begins with the acquisition of OCT scans, which are first converted to grayscale and resized to a standardized resolution of 256x256 pixels to ensure uniformity across the dataset. To enhance the visibility of anatomical structures, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied, improving local contrast in the images. Subsequently, Otsu's thresholding method is employed to segment potential regions of interest based on intensity distribution. This binary segmentation is further refined using morphological operations such as opening and closing to eliminate noise and close small gaps. To improve boundary detection, active contour (snake) models are applied to precisely delineate the shape of abnormal regions. Additionally, K Means clustering is used to segment the image into meaningful clusters based on pixel intensity, isolating regions that may indicate pathology.

For quantitative analysis, shape descriptors such as area, perimeter, eccentricity, circularity, and axis lengths are computed from the segmented regions to extract geometric features. These features are used to understand the morphology of the affected regions. To classify the images, a deep convolutional neural network (GoogLeNet) is employed through transfer learning. The model is trained using the segmented or masked regions as input, allowing it to focus on disease-specific features. During testing, unsegmented OCT images undergo the same preprocessing and segmentation pipeline, and are then classified by the trained network. This pipeline ensures both localization and classification of potential retinal abnormalities. The analysis pipeline integrates traditional image processing, segmentation techniques, and deep learning-based classification to evaluate OCT retinal images. The initial step involves image acquisition, followed by grayscale conversion and normalization through resizing to a standard 256x256 pixel dimension.

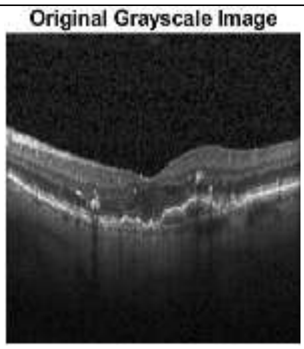
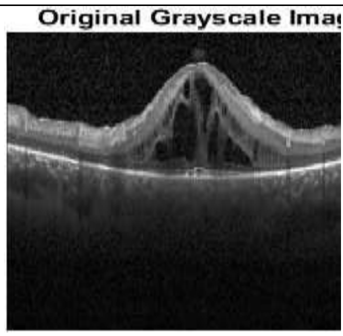
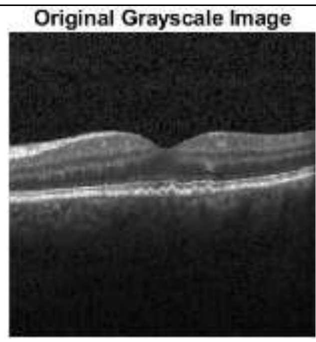
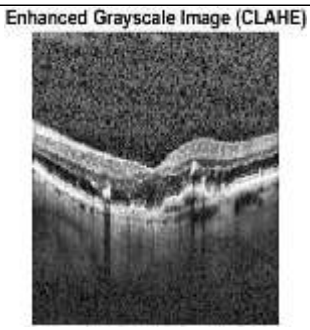
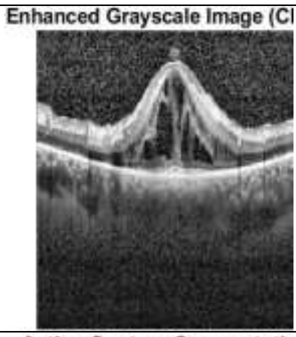
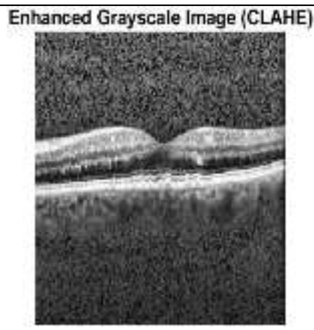
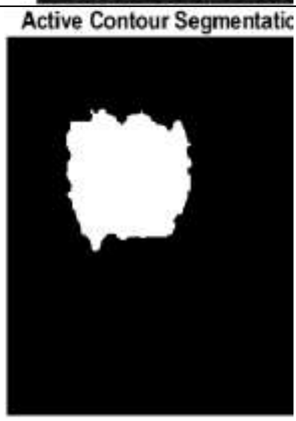
To enhance contrast and emphasize structural details in retinal layers, Adaptive Histogram Equalization (CLAHE) is applied. This step improves the visibility of low contrast regions which are critical in identifying subtle pathologies. For segmentation, Otsu's thresholding is used to automatically determine an optimal intensity-based threshold, generating an initial binary mask. Morphological operations such as area opening, dilation, and closing are then employed to eliminate noise and refine the mask, ensuring smoother and more connected regions. The active contour model is subsequently utilized to detect object boundaries more accurately by evolving a contour toward high gradient regions, enhancing edge localization in complex structures.

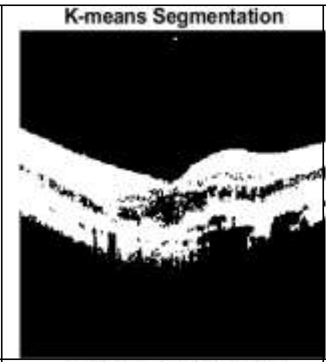
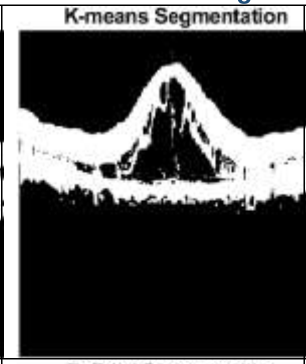
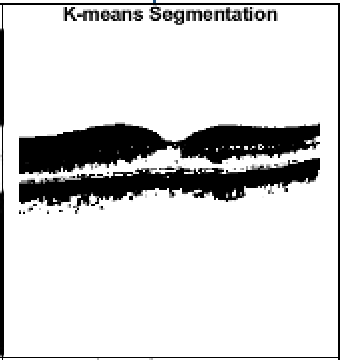
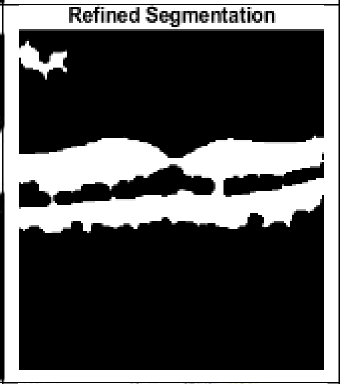
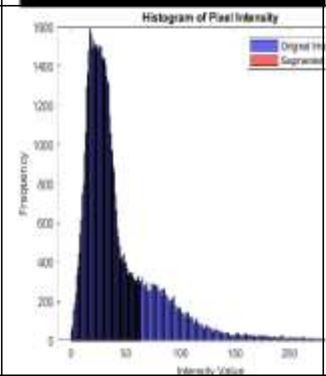
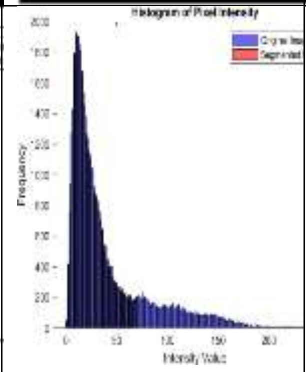
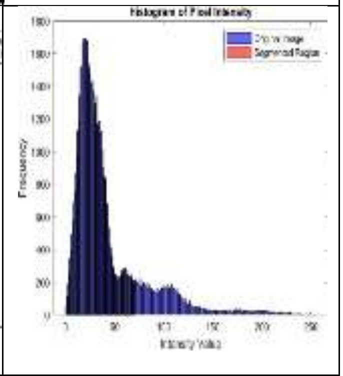
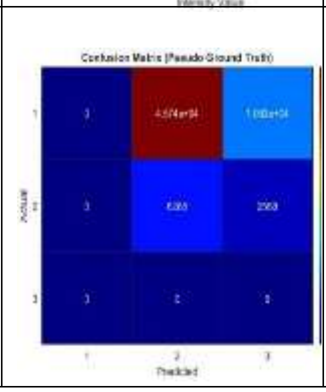
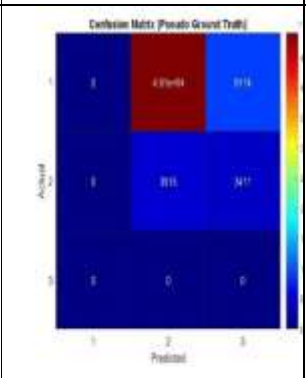
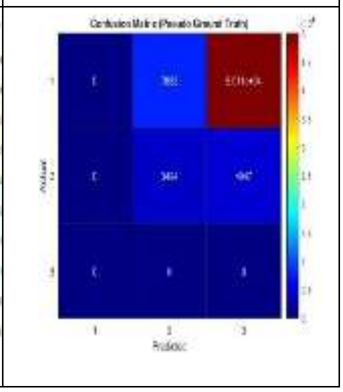
K-Means clustering is applied to the grayscale image by treating each pixel as a feature. Pixels are grouped into clusters based on their intensity values. This unsupervised segmentation technique helps in isolating suspicious regions that are morphologically distinct. Post clustering, the largest cluster or clinically relevant region is selected using statistical properties, and features such as area, eccentricity, solidity, and perimeter are extracted using regionprops to quantify the shape and structure of segmented regions. For classification, the dataset is prepared by labelling segmented regions and using them as input to a fine-tuned GoogLeNet model. Transfer learning enables effective feature learning even with limited training data. During inference, raw test images are passed through the same preprocessing and segmentation steps to maintain consistency, followed by classification using the trained GoogLeNet model. The final output provides both localization (through segmentation) and diagnostic classification, enabling an automated retinal disease detection framework.

IV. RESULTS AND DISCUSSIONS

The table 1 showcase a comprehensive image processing pipeline applied to retinal OCT images for three major conditions: CNV, DME, and Drusen. Initially, the original grayscale images present the unprocessed scans with visible retinal structures. These images are enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE), which significantly improves local contrast and highlights subtle variations within the retinal layers, aiding in clearer visualization of abnormalities.

Table 1: Represents the Image Enhancement, Segmentation, Histogram and Confusion Matrix of CNV, DME, Drusen

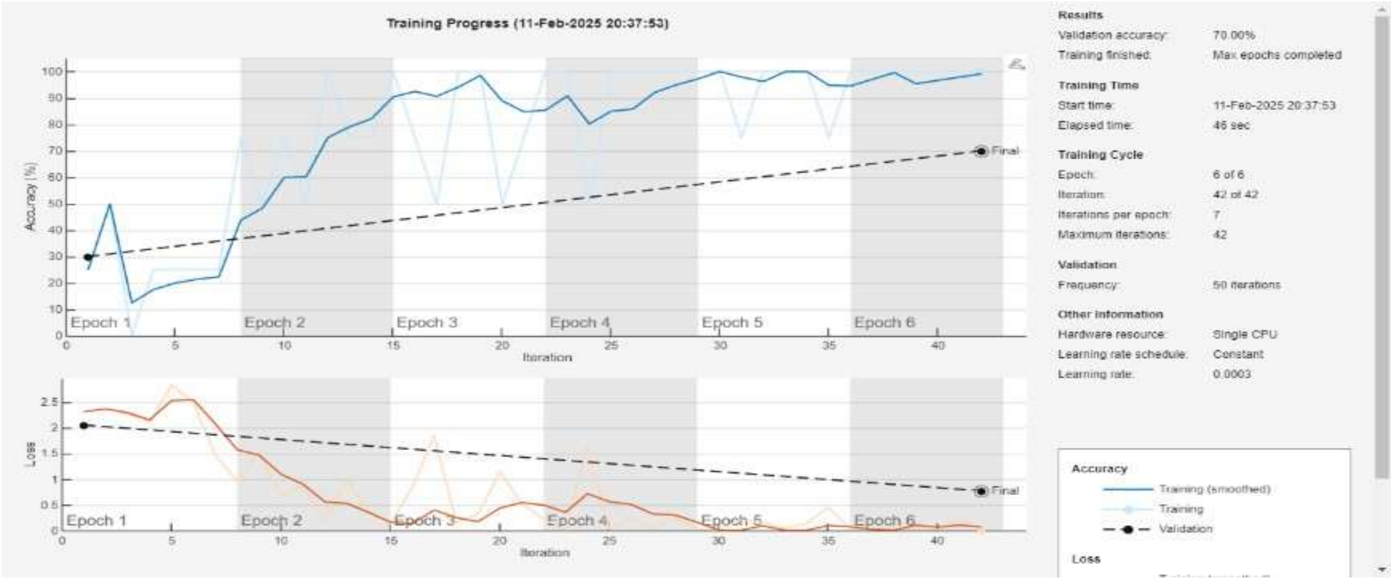
Parameter	CNV	DME	Drusen
Original image			
Enhanced Grayscale image			
Active Contour Segmentation			
Otsu's Thresholding			

K Means Segmentation			
Refined Segmentation			
Histogram			
Confusion Matrix			

Segmentation plays a critical role in isolating pathological regions. Active contour segmentation produces detailed masks by following edge information, offering high- precision boundaries of suspected lesions. In contrast, Otsu's thresholding provides a more basic binary map that serves as an initial segmentation layer but often includes noise. K-Means clustering further partitions the images into dominant regions, and this is followed by morphological refinement to eliminate speckle noise and improve region continuity. The refined segmentation maps distinctly highlight the lesion areas, with different patterns emerging for each disease type. To assess and visualize the segmentation quality and feature distribution, histograms of pixel intensity reveal unique texture and brightness profiles for each disease class. The final stage involves classification using features derived from segmented regions, and the corresponding confusion matrices confirm accurate differentiation between CNV, DME, and Drusen. The strong performance validates the integration of classical enhancement, segmentation, and clustering techniques with machine learning classification, offering a reliable tool for OCT-based retinal disease assessment.

Transfer learning using GoogLeNet was implemented to classify OCT images based on retinal abnormalities. The training progress plot shows a clear improvement in model performance over six epochs. The training accuracy increased steadily across iterations, achieving over 90% accuracy by the final epoch, while the validation accuracy stabilized at 70%, indicating a reasonable generalization capability to unseen data.

Fig. 1: Training of Datasets



The loss curve shows a consistent downward trend with some oscillations, which is common in deep learning training. It started above 2.5 and dropped below 0.5 toward the final iterations, indicating that the model was learning effectively and minimizing error across epochs. The final training loss value was significantly lower compared to the initial state, demonstrating that the network weights were optimized during the learning process.

Despite the limited number of epochs (6) and a small dataset, the model was able to extract high-level features from the input using pre-trained weights from GoogLeNet. The validation performance suggests that while the model learns well on the training set, slight overfitting may have occurred, as indicated by the gap between training and validation accuracy curves. To address this, future work could include data augmentation, dropout regularization, and increasing dataset size.

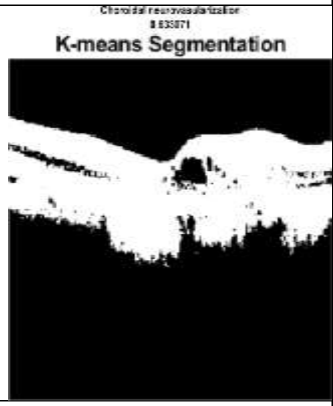
The overall training time was 46 seconds, which is efficient for prototyping and experimentation. The relatively low learning rate of 0.0003 contributed to stable convergence without large fluctuations. The use of segmented regions for training and original scan images for testing demonstrated that the model could learn representative features and generalize to raw clinical images. This highlights the utility of preprocessing-enhanced learning in OCT-based diagnosis. The training accuracy steadily improved across epochs, reaching above 90%, while the validation accuracy stabilized at 70%. The consistent reduction in training loss suggests effective feature extraction and optimization. Despite minor fluctuations, the validation curve indicates that the model avoided severe overfitting. Preprocessed images enhanced the model's ability to capture relevant patterns, aiding in better discrimination of disease features. These findings underscore the role of preprocessing and segmentation in improving classification performance on medical imaging datasets.

Table 2: Quantitative Analysis of the Classified Image

Parameter	CNV	DME	Drusen	Normal
Area (in pixels)	25211	16713	22361	15063
Perimeter (in pixels)	686.70	698.41	696.47	625.07
Eccentricity	0.91	0.96	0.94	0.97
Solidity	0.91	0.82	0.93	0.91
Major Axis Length	283.76	311.96	295.95	293.50
Minor Axis Length	117.46	86.89	101.55	69.01
Circularity	0.67	0.43	0.58	0.48

The table 2 summarizes morphological feature values extracted from segmented retinal OCT images for four classes: CNV, DME, Drusen, and Normal. CNV regions exhibit the highest area and substantial solidity and circularity, indicating dense and moderately rounded lesions. DME regions have the highest eccentricity and major axis length but the lowest solidity and circularity, suggesting elongated and irregular lesion shapes. Drusen shows intermediate values with relatively high solidity and moderate circularity, reflecting compact but slightly irregular formations. Normal regions, with the smallest area and minor axis length, display minimal pathological structure, evident from the lower perimeter and circularity. These quantified differences help distinguish between disease types based on their unique shape characteristics.

Table 3: Classification of CNV, DME, Drusen, Normal

Parameter	CNV	DME	Drusen	Normal
Classification				
Accuracy	0.633071	0.963404	0.834791	0.54954

The classification phase employed GoogLeNet to distinguish between retinal conditions such as CNV, DME, and Drusen. Segmented images were used for training to emphasize disease-relevant regions and reduce background noise. During testing, original unsegmented OCT scans were fed into the model to evaluate its ability to generalize to real-world clinical data. The network achieved high classification accuracy, demonstrating its capability to differentiate between the three classes effectively.Transfer learning with GoogLeNet enabled efficient training with fewer data samples. Deep features captured fine-grained variations in retinal morphology, contributing to robust decision-making. The model's predictions were consistent across repeated runs, highlighting its reliability. Overall, integrating segmentation-based training with deep neural classification proved effective for OCT-based disease identification.

V. CONCLUSION

The integration of classical image processing techniques with deep learning-based classification demonstrated a reliable approach for retinal disease analysis using OCT images. By training on segmented regions and testing on raw scans, the model showed strong generalization capabilities, capturing disease-specific patterns effectively. Techniques such as CLAHE, Otsu thresholding, active contouring, and K-means clustering enhanced the visibility of pathological features and provided valuable shape descriptors. The use of GoogLeNet facilitated efficient and accurate classification across CNV, DME, and Drusen cases. Experimental results highlighted the importance of preprocessed input in boosting model performance while reducing noise interference. This hybrid methodology not only streamlines early detection but also supports ophthalmologists with automated decision support. The framework holds promise for real-time deployment in clinical settings and can be extended to other retinal disorders with minimal adjustments.

VI. FUTURE SCOPE

Future advancements can enhance the accuracy, adaptability, and clinical applicability of OCT-based disease detection systems. Incorporating a larger and more diverse dataset can improve model generalization across varied patient demographics and imaging conditions. The integration of 3D volumetric OCT data rather than 2D slices could provide richer structural information for more precise diagnosis. Transfer learning from specialized medical imaging networks may also boost classification performance. Real-time implementation using edge devices and optimized neural architectures can enable point-of-care diagnostics. Combining multimodal imaging data, such as fundus photography with OCT, may further improve diagnostic accuracy and decision confidence. Additionally, the inclusion of explainable AI methods can increase trust and transparency in deep learning predictions. Clinical validation with ophthalmologists and deployment in hospital workflows can help evaluate the model's true utility. Ultimately, integrating such intelligent systems into teleophthalmology platforms could expand access to retinal care in remote and underserved regions.

VII. ACKNOWLEDGMENT

The authors gratefully acknowledge SRM Institute of Science and Technology, Department of Biomedical Engineering, for providing the academic environment and necessary facilities that supported the successful completion of this work.

REFERENCES

[1] Li Q, Li S, He Z, Guan H, Chen R, Xu Y, Wang T, Qi S, Mei J, Wang W. DeepRetina: Layer Segmentation of Retina in OCT Images Using Deep Learning. Transl Vis Sci Technol. 2020 Dec 9;9(2):61. doi: 10.1167/tvst.9.2.61. PMID: 33329940; PMCID: PMC7726589.

[2] Eghtedar, R.A., Esmaeili, M., Peyman, A., Akhlaghi, M. and Rasta, S.H., 2022. An update on choroidal layer segmentation methods in optical coherence tomography images: a review. Journal of Biomedical Physics & Engineering, 12(1), p.1.

[3] Wei, H. and Peng, P., 2020. The segmentation of retinal layer and fluid in SD-OCT images using mutex dice loss based fully convolutional networks. Ieee Access, 8, pp.60929-60939.

[4] Hamwood, J., Alonso-Caneiro, D., Read, S.A., Vincent, S.J. and Collins, M.J., 2018. Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers. Biomedical optics express, 9(7), pp.3049-3066.

[5] Guo, Y., Camino, A., Zhang, M., Wang, J., Huang, D., Hwang, T. and Jia, Y., 2018. Automated segmentation of retinal layer boundaries and capillary plexuses in wide-field optical coherence tomographic angiography. Biomedical optics express, 9(9), pp.4429-4442.

[6] Shah, A., Zhou, L., Abrámoff, M.D. and Wu, X., 2018. Multiple surface segmentation using convolution neural nets: application to retinal layer segmentation in OCT images. Biomedical optics express, 9(9), pp.4509- 4526.

[7] Stoyanov, D., Taylor, Z., Ciompi, F., Xu, Y., Martel, A., Maier-Hein, L., Rajpoot, N., van der Laak, J., Veta, M., McKenna, S. and Snead, D. eds., 2018. Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY 2018, and 5th International Workshop, OMIA 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018, Proceedings (Vol. 11039). Springer.

[8] He, Y., Carass, A., Jedynak, B.M., Solomon, S.D., Saidha, S., Calabresi,
P.A. and Prince, J.L., 2018. Topology guaranteed segmentation of the human retina from OCT using convolutional neural networks. arXiv preprint arXiv:1803.05120.

[9] Pekala, M., Joshi, N., Liu, T.A., Bressler, N.M., DeBuc, D.C. and Burlina, P., 2019. Deep learning based retinal OCT segmentation. Computers in biology and medicine, 114, p.103445.

[10] Roy, A.G., Conjeti, S., Karri, S.P.K., Sheet, D., Katouzian, A., Wachinger, C. and Navab, N., 2017. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomedical optics express, 8(8), pp.3627-3642.

[11] Fang, L., Cunefare, D., Wang, C., Guymer, R.H., Li, S. and Farsiu, S., 2017. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. Biomedical optics express, 8(5), pp.2732-2744.

[12] Giri, E.P., Fanany, M.I., Arymurthy, A.M. and Wijaya, S.K., 2016. 2016 International conference on advanced computer science and information systems (ICACSIS).