Retinal Layer Segmentation and Classification in OCT Images for Disease Detection

Maria V. Leyba Mesa, Elijah Ray, Bayan Ahmad, *Member, IEEE,* Buket D. Barkana, *Member, IEEE*

***Abstract*— Optical coherence tomography (OCT) image analysis plays a crucial role in detecting and monitoring retinal diseases, such as diabetic macular edema and age-related macular degeneration. In our work, we introduce a retinal layer segmentation algorithm for OCT images, organizing the layers into four primary groups. To enhance segmentation accuracy, we designed new preprocessing techniques including background adjustment, adaptive denoising, and rotational correction. An interdependent segmentation model was developed using Otsu's method, edge detection, and morphological operators. To evaluate the effectiveness of the segmented layers, we extracted statistical and texture-based features and applied KNN, SVM, Naïve Bayes, and Neural Network classifiers on a dataset of 84,495 OCT images. The classifiers demonstrated strong performance, validating the approach.**

I. Introduction

Approximately 288 million people worldwide are projected to be affected with age-related macular degeneration (AMD) by the year 2040 [1] and 161 million people with diabetic macular edema (DME) by the year 2045 [2]. AMD and DME are two of the most prominent retinal diseases. The retina, known as the innermost layer of the eye, plays a large role in facilitating the processing of light signals into electrical signals [3]. Thus, damage to retinal tissue causes a mistranslation between the two signal types and hinders the brain's perception. This misperception directly correlates to the visual impairment experienced by those with retinal diseases. Retinal diseases are often assessed through optical coherence tomography (OCT), which allows ophthalmologists to view a cross-section of the retina [4] to determine disease type and progression and advise treatments. With the projected increase in AMD and DME cases, developing OCT image segmentation and analysis methods is becoming more important.

The retina comprises ten layers, some with vastly different contents and functions (See Fig. 1). The inner limiting membrane (ILM) is the innermost subretinal layer that separates the retina from the vitreous humor. This is followed by the nerve fiber layer (NFL). Ganglion cells are housed in the following layer, the ganglion cell layer. A series of subsequent plexiform and nuclear layers comprise much of the retinal structure. The rod and cone cell bases are in the external limiting membrane (ELM) [3], which also separates the previous layers from the photoreceptor segments and retinal pigment epithelium (RPE) [4]. The posterior RPE also comprises some of the Bruch's membrane complex [5]. The choroid is responsible for providing blood supply to the outermost layers of the retina [3]. The normal aging of the eye is reflected in OCT scans. However, these changes are presented differently than in AMD- and DME-conditioned eyes. In normal OCT images of young individuals, the subretinal layers alternate with hyperreflective and hyporeflective bands that are clearly defined with prominent boundaries. The bands are continuous, and outer subretinal layers exhibit a thin and smooth structure. As healthy aging progresses, the ELM and outer bands become thicker and display more irregularities. Additionally, band visibility and definition are reduced [6].

AMD is understood to be a result of inflammatory responses to oxidative stress, which is associated with the aging process. It is associated with drusen depositions, which are aggregates of extracellular compositions. There are two branches of AMD, dry and wet. Dry AMD is more prevalent and characterized by a thicker accumulation of drusen, along with a more prominent Bruch's membrane and a loss of photoreceptors. Choroidal neovascularization (CNV) is a critical aspect of wet AMD and is a response to subretinal cell migration [7]. Diabetic Retinopathy (DR) is a result of having diabetes mellitus condition. Diabetic macular edema (DME) presents in DR as fluid pockets that develop in the neural retina, causing it to thicken [7].

The paper is organized as follows: Section two presents the preprocessing and four-layer segmentation methodology, section three shows the analysis of the segmented layers for each group, and section four summarizes the results. Fig.2 shows a flow diagram of the proposed methodology.

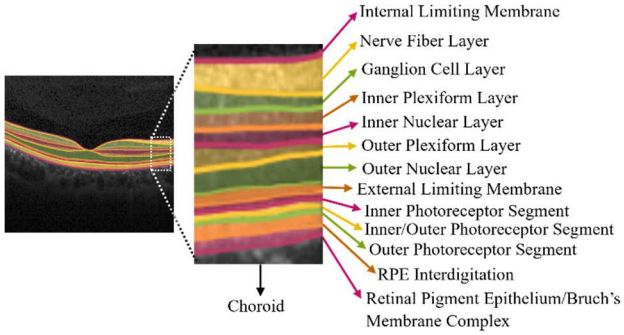

Figure 1. Retinal layers annotated on OCT image.

M. V. Leyba Mesa is with the Biomedical Engineering Department, The University of Akron, Akron, OH, 44325 USA (phone: 216-250-6850; e-mail: mvl16@uakron.edu).

E. Ray is with the Biomedical Engineering Department, The University of Akron, Akron, OH, 44325 USA (e-mail: emr115@uakron.edu).

B. Ahmad is with the Biomedical Engineering Department, The University of Akron, Akron, OH, 44325 USA (e-mail: ba82@uakron.edu).

B. D. Barkana is with the Biomedical Engineering Department, The University of Akron, Akron, OH, 44325 USA (e-mail: bbarkana@uakron.edu).

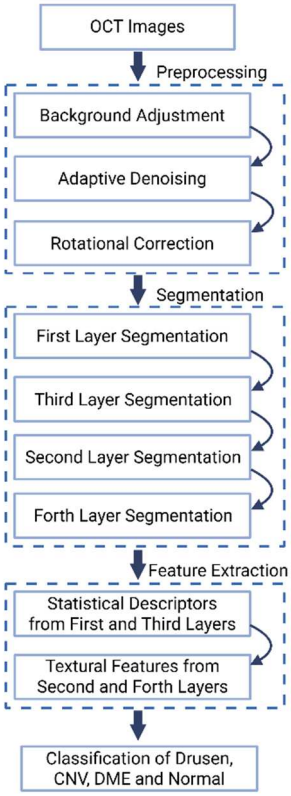

Figure 2. Architecture of the proposed methodology.

II. METHOD

In this study, we present a methodology for segmenting retinal layers into four layers: layer one includes the ILM and the NFL, the second layer includes all layers between the NFL and the ELM, and the third layer is composed of the photoreceptor segments and the remaining regions of the retina, with the choroid composing the fourth layer.

The dataset contains 84,495 OCT images (JPEG) and subfolders for four categories (NORMAL, CNV, DME, DRUSEN) [8]. The Normal category includes 27323 images, the CNV category includes 37206 images, the DME category includes 11349 images and the DRUSEN category includes 8617 images. Fig.3 shows the four categories. The OCT images underwent both preprocessing and segmentation processes to obtain the four layers previously mentioned.

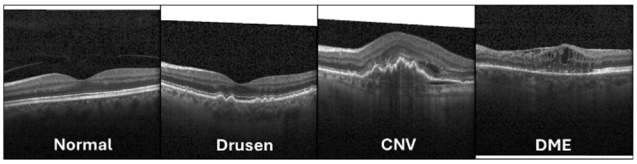

Figure 3. Representative OCT images for Normal, Drusen, CNV, and DME categories.

A. Preprocessing

The preprocessing stage includes background adjustment, adaptive denoising, and rotational correction to enhance the raw OCT images and prepare them for the segmentation stage.

Background adjustment: The background of some images in the OCT dataset has a white region located at the outer parts of the image (See Fig.3), which affects the segmentation process. The first step in the preprocessing stage is to remove the unwanted regions from the background. An edge detection algorithm was used to recognize such regions and convert them into black pixels.

Adaptive denoising: The dataset contains some images with a significantly low signal-to-noise ratio. We employed a histogram-based thresholding approach to filter noise. An adaptive thresholding value was determined based on the noise level of any one image. A K×L block of the background with distributed noise, I_n, was used, where K and L were set to 50. The block's average pixel intensity value, μ_n, was calculated to define an adaptive threshold value to filter noise. Assuming the pixel intensity values of I_n have a relatively normal distribution, approximately half of the pixels in I_n should have an intensity value below μn, due to the shape of a cumulative distribution curve and the position of μn. For any pixel in I_n the univariate normal pdf [9] can be defined as

$$P[0 < X \leq Th] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{Th} e^{-0.5\left(\frac{x-\mu_n}{\sigma}\right)^2} dx. \quad (1)$$

Any pixel intensity, x, obeys the normal probability law with mean μ_n and standard deviation σ. A threshold value, Th, was determined to be at the value of $\mu_n + \sigma$ to eliminate the background while keeping the integrity of the region of interest (ROI) edges. It corresponded to approximately 85 % of pixels within I_n.

Algorithm 1: Adaptive denoising Filter

1: **Given:** Original Image I_n, Adjusted Background Output I_{ADJ}.
2: Create a $K \times L$ pixel block, $b_{K\times L}$. Where K=L=50.
3: Place $b_{K\times L}$ on the center of I_n.
4: Ensure $b_{K\times L}$ does not include any white pixels.
5: Determine the average pixel intensity, μ_n in $b_{K\times L}$.
6: Calculate the standard deviation, σ.
7: Calculate the threshold value, Th.
8: On I_{ADJ} turn all pixels of values above Th to 0.

Rotational correction: Due to the nature of OCT image acquisition, some images are tilted. To ensure proper alignment, a custom image rotation method was developed. The method detects any tilt in the images and performs the necessary rotation to align the image horizontally. Fig.4a shows an image with a tilted ROI. This step improves the segmentation of the four layers.

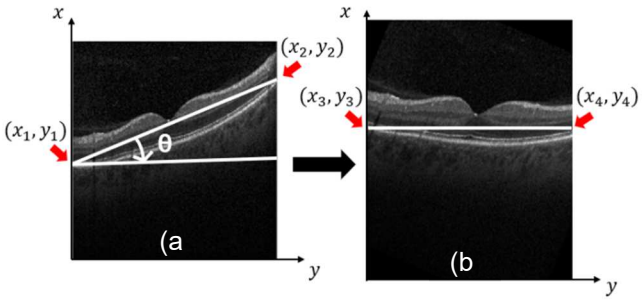

Figure 4. OCT image rotation. A) pre- and b) post- rotation.

The rotation process first detects the coordinates of the photoreceptor segment; a simple thresholding method is sufficient since the segment has the highest image intensity. The rotation angle required to align the image was calculated using the arctangent function (2).

$$\theta = \tan^{-1}\frac{\Delta y}{\Delta x} = \tan^{-1}\frac{(y_2 - y_1)}{(x_2 - x_1)} \quad (2)$$

Δy is the vertical distance and Δx is the horizontal distance. The angle θ was then applied to align the detected edge points horizontally, thus correcting any tilt. The rotation ensures that all images are consistently aligned, minimizing errors caused by tilt and improving the accuracy of subsequent processing steps. (x_3, y_3) and (x_4, y_4) become the new coordinates of the rotated image as . (Fig. 4b)

Algorithm 2: Rotational Correction

1: **Given:** Segmented Image I_{seg}.
2: Find the first non-black pixel in the first column, $I_{seg(x_1,y_1)}$.
3: Find the first non-black pixel in the last column, $I_{seg(x_2,y_2)}$.
4: Calculate θ.
5: Rotate the image θº.

B. Segmentation

Four-layer groups were segmented using a series of stages: binarization, edge detection, morphological operations, and adjacent pixel clustering.

First and Third Layer Segmentation: The segmentation of the first layer includes Otsu's method, edge detection, and morphological closing. Otsu's method was used for binarization. It maximizes the variance between the background and region of interest (ROI), ensuring that the threshold is selected so that the pixels in each region are as homogeneous as possible. The method calculates the weighted sum of variances for all possible threshold values and chooses the one that minimizes the within-class variance, effectively distinguishing the ROI from the background without needing user-defined parameters. Canny edge detection was applied to the binarized images. In a multi-stage algorithm, the Canny filter first applies a Gaussian filter to the image to reduce noise, ensuring that minor variations in pixel intensity do not interfere with edge detection. Next, the algorithm calculates the intensity gradients of the image by applying Sobel filters in both horizontal and vertical directions, highlighting areas of rapid intensity change, which are potential edges. Non-maximum suppression is applied to sharpen these edges, keeping only the local maxima in the gradient direction and thinning the edges. Then, two thresholds (high and low) are applied using a technique known as hysteresis thresholding. Pixels with gradient values above the high threshold are marked as edges, while those between the two thresholds are only considered edges if they are connected to a pixel above the high threshold.

Morphological closing was used to smooth edges and close small gaps, holes, or breaks within an edge. It consists of two fundamental operations in which dilation is followed by erosion. Combining these operations helps retain the object's overall size and shape while smoothing irregularities.

We used "disk' as the structuring element with a radius of 5. Fig.5a shows the first layer segmentation results. The first layer of segmentation is followed by the third layer.

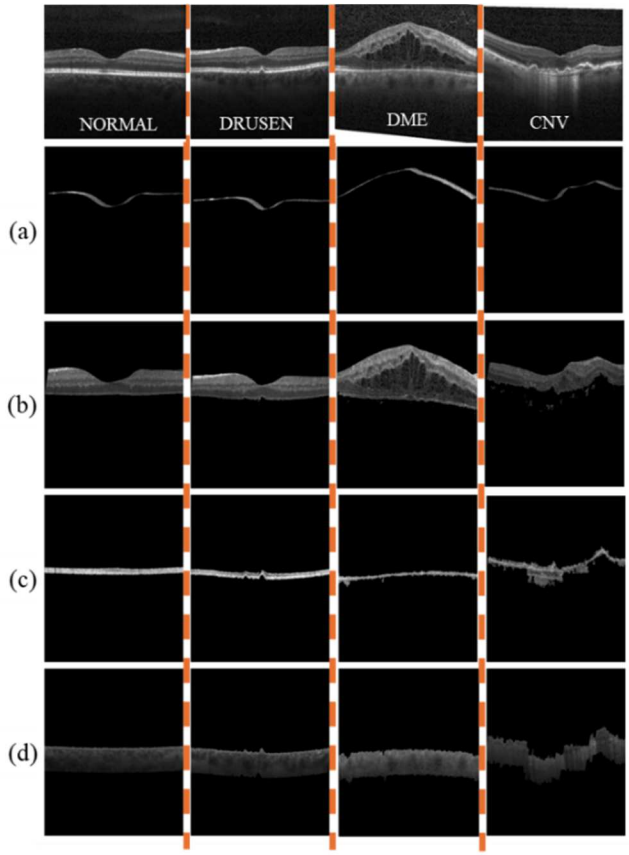

Figure 5. Layer segmentation results of four categories.

The clusters of adjacent pixels in the binarized image were detected. The photoreceptor segments create the third layer in this work, and these segments have the highest intensity values in the OCT images. Consequently, the associated pixel

cluster is usually larger than the others, and the third layer lies below the first layer. Our approach takes advantage of these characteristics. However, the algorithm does not assume that the largest cluster belongs to the third layer and checks the segmented first layer to ensure the third layer's location. The morphological closing operation closed small gaps and breaks within the cluster. Fig.5c shows the third layer segmentation results.

Second Layer Segmentation: The second layer lies between the top and third layers. The binary representations of the top and third layers were used as boundaries of the second layer. Fig.5b shows the second layer segmentation results.

Fourth Layer Segmentation: A region beneath the third layer was segmented into the fourth layer. The approach allowed for the extraction of the entire retinal structure, including the space below the retinal layers, ensuring a complete representation of the OCT image. Fig.5d shows the segmentation results of the fourth layer.

III. Analysis

We assessed the effectiveness of four segmented layers in describing the OCT categories. A set of features was extracted from each layer, and a one-way ANOVA was used to analyze them. Table 1 describes the extracted features. Since some first- and third- layers' structures show signs of AMD and DME, such as discontinuity and irregularities, we used low moment statistical descriptors -mean, variance, skewness, and kurtosis – of the layers and the first derivative of the layers' upper boundaries. In the presence of DME the third layer may exhibit thickening and fluid pockets. Texture features - contrast, correlation, energy, and homogeneity – were used to represent the changes in this layer using the Gray-Level Co-Occurrence Matrix to quantify spatial relationships between pixel intensities in an image [10]. Contrast captures the local variations in intensity and highlights structural changes in the layer. Correlation reflects the linear dependency of gray levels in adjacent pixels. It may obtain fluid packets patterns. Energy measures the uniformity and indicates how much of the region is dominated by repetitive patterns like abnormal structures in the layer as homogeneity indicates how uniform the texture is.

Z-score standardization was performed to normalize the data ensuring all features are on the same scale before analyzing the data. The Z-score, Z, was calculated (3).

$$Z = \frac{x - \mu}{\sigma} \tag{3}$$

The mean value, μ, and the standard deviation, σ, was calculated for all datapoints, x. Creating a new data set with a new μ value of 0 and a σ value of 1.

TABLE I. Feature Extraction methods

Layer	Features
First Layer	Statistical descriptors of the layer and its first-degree derivative
Second Layer	Textural features: contrast, correlation, energy, and homogeneity [10]
Third Layer	Statistical descriptors of the layer and its first-degree derivative
Fourth Layer	Textural features: contrast, correlation, energy, and homogeneity [10]

Furthermore, fundamental machine learning models, K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Naïve Bayes (NB), and Neural Networks (NN) were designed to evaluate the segmented layers' performance in differentiating four categories. The hyperparameters of the models are given in Table 2.

TABLE II. KNN, SVM, Naïve Bayes, and NN hyperparameters.

Model	Hyperparameters
KNN	Cubic Kernel 10 Neighbors Distance Metric: Minkowski (cubic) Distance Weight: Equal
SVM	Cubic Kernel Automatic Kernel Scale Box Constraint Level: 1 Multiclass Coding: One-vs-One
NB	Gaussian Naïve Bayes
NN	Medium Neural Network 1 Fully Connected Layer First Layer Size: 25 Activation: ReLU

Performance metrics of the models are given in Table 3.

TABLE III. KNN, SVM, Naïve Bayes, and NN overall performace metrics.

Model	Acc	Se	Sp	AUC
KNN	0.8702	0.8236	0.9510	0.9708
SVM	0.9787	0.9748	0.9927	0.9948
NB	0.7286	0.6777	0.9005	0.8759
NN	0.9817	0.9788	0.9938	0.9985

The models demonstrated strong performance, indicating that the segmented layers effectively represent abnormalities in OCT images. The Naïve Bayes model, a probabilistic approach, performed poorly compared to the others due to its assumption that all features are independent. In reality, the segmented layers and features in the dataset are interdependent, which affected the model's performance. The Neural Network Model performed the best out of the four models. The convergence information of the trained neural network included 1000 iterations, a training loss of 0.013, and a step of 0.01. The proposed model's performance is compared with previous works in Table 4.

TABLE IV. COMPARISON OF THE PROPOSED MODEL TO PREVIOUS WORKS.

	Acc	**Se**	**Sp**	**AUC**
[11]	0.9550	0.9465	0.9854	-
[12]	0.9733	0.9467	**1.0000**	0.9733
[13]	0.9286	0.9237	0.9340	0.9783
This work (NN)	**0.9817**	**0.9788**	0.9938	**0.9985**

As shown in the table, our proposed approach performs with higher accuracy, sensitivity and AUC compared to previous works.

IV. CONCLUSION

We developed a segmentation model to categorize retinal layers into four groups. To evaluate the model's effectiveness, a dataset of 84,495 OCT images across four categories (NORMAL, CNV, DME, DRUSEN) was analyzed using statistical and texture-based features. Four machine learning models demonstrated strong performance, validating the robustness of the segmentation model. Our proposed segmentation demonstrates potential for the development of a computer-aided diagnosis system for timely diagnosis and treatment of retinal diseases.

REFERENCES

[1] G. Kaur and N.K. Singh, “Inflammation and retinal degenerative diseases,” Neural Regeneration Research, vol. 18, no. 3, pp. 513-518, 2023.

[2] T.T. Wong and T.E. Tan, “The Diabetic Retinopathy “Pandemic” and Evolving Global Strategies: The 2023 Friedenwald Lecture,” Invest. Ophthalmol. Vis. Sci., vol. 64, no. 15, pp. 47, 2023.

[3] N. Mahabadi and Y. Al Khalili, “Neuroanatomy, Retina,” StatPearls Publishing, 2024, [online] Available: https://www.ncbi.nlm.nih.gov/books/NBK545310/ Accessed on September 3, 2024

[4] D. Fu, H. Tong, S. Zheng, L. Luo, F. Gao, and J. Minar, “Retinal status analysis method based on feature extraction and quantitative grading in OCT images,” BioMed Eng OnLine, vol. 15, no. 87, 2016.

[5] E. Borrelli, D. Sarraf, K.B. Freund, and S.R. Sadda, “OCT angiography and evaluation of the choroid and choroidal vascular disorders,” Progress in retinal and eye research, vol. 67, pp. 30-55, 2018.

[6] S. Chen, O. Abu-Omar, D. Kar, J.D. Messinger, Y. Hwang, E.M. Moult, et al., “Ultrahigh resolution OCT markers of normal aging and early age-related macular degeneration,” Ophthalmology Science, vol. 3, no. 3, pp. 100277, 2023.

[7] G. Kaur and N.K. Singh, “Inflammation and retinal degenerative disease,” Neural Regenerative Research, vol.18, no. 3, pp. 513-518, 2023.

[8] D. Kermany, K. Zhang, and M. Goldbaum, “Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification,” Mendeley Data v2, vol.2, no. 2, 2018.

[9] H. Stark and J.W. Woods, “Probability, Random Process, and Estimation Theory for Engineers”, Prentice Hall, pp. 44-45, 1986.

[10] Y. Park and J.M. Guldmann, “Measuring Continuous Landscape Patterns with Gray-Level Co-Occurrence Matrix (GLCM) Indices: An Alternative to Patch Metrics?” Ecological Indicators, pp.1-18, 2020.

[11] V. Das, S. Dandapat, P. Kumar Bora,
“Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images”, Biomedical Signal Processing and Control, vol. 54, 2019.

[12] Hussain MA, Bhuiyan A, D. Luu C, Theodore Smith R, H. Guymer R, Ishikawa H, et al. “Classification of healthy and diseased retina using SD-OCT imaging and Random Forest algorithm”. PLoS ONE, 2018

[13] Y. Rong et al., "Surrogate-Assisted Retinal OCT Image Classification Based on Convolutional Neural Networks," in *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 253-263, Jan. 2019.