

# TUNABLE WAVELET UNIT BASED CONVOLUTIONAL NEURAL NETWORK IN OPTICAL COHERENCE TOMOGRAPHY ANALYSIS ENHANCEMENT FOR CLASSIFYING TYPE OF EPIRETINAL MEMBRANE SURGERY

An Le<sup>1</sup>, Nehal Mehta<sup>2</sup>, William Freeman<sup>2</sup>, Ines Nagel<sup>2</sup>, Melanie Tran<sup>2</sup>, Anna Heinke<sup>2</sup>,  
Akshay Agnihotri<sup>2</sup>, Lingyun Cheng<sup>2</sup>, Dirk-Uwe Bartsch<sup>2</sup>, Hung Nguyen<sup>1</sup>, Truong Nguyen<sup>1</sup>, Cheolhong An<sup>1</sup>  
<sup>1</sup>Electrical and Computer Engineering Department, University of California San Diego, La Jolla, CA 92093, USA  
{dole,hun004,tqn001,chan}@ucsd.edu

<sup>2</sup>Jacobs Retina Center, Shiley Eye Institute, University of California San Diego, La Jolla, CA 92093, USA  
{nnmehta,wrfreeman,inagel,mdtran,aheinke,aagnihotri,lcheng,dbartsch}@health.ucsd.edu

**Abstract**—In this study, we developed deep learning-based method to classify the type of surgery performed for epiretinal membrane (ERM) removal—either internal limiting membrane (ILM) removal or ERM-alone removal. Our model, based on the ResNet18 convolutional neural network (CNN) architecture, utilizes postoperative optical coherence tomography (OCT) center scans as inputs. We evaluated the model using both original scans and scans preprocessed with energy crop and wavelet denoising, achieving 72% accuracy on preprocessed inputs, outperforming the 66% accuracy achieved on original scans. To further improve accuracy, we integrated tunable wavelet units with two key adaptations: Orthogonal Lattice-based Wavelet Units (OrthLatt-UwU) and Perfect Reconstruction Relaxation-based Wavelet Units (PR-Relax-UwU). These units allowed the model to automatically adjust filter coefficients during training and were incorporated into downsampling, stride-two convolution, and pooling layers, enhancing its ability to distinguish between ERM-ILM removal and ERM-alone removal, with OrthLatt-UwU boosting accuracy to 76% and PR-Relax-UwU increasing performance to 78%. Performance comparisons showed that our AI model outperformed a trained human grader, who achieved only 50% accuracy in classifying the removal surgery types from postoperative OCT scans. These findings highlight the potential of CNN based models to improve clinical decision-making by providing more accurate and reliable classifications. To the best of our knowledge, this is the first work to employ tunable wavelets for classifying different types of ERM removal surgery.

**Index Terms**—Biomedical image processing, Retinal images, Machine learning, Convolutional neural networks, Residual neural networks, Discrete wavelet transforms.

This article has been supported in part by UCSD Vision Research Center Core Grant P30EY022589, NIH grant R01EY016323, NIH Grant R01EY033847, National Institutes of Health Bridge2AI Project (OT2OD032644), an unrestricted grant from Research to Prevent Blindness, NY, and unrestricted funds from UCSD Jacobs Retina Center.

## I. INTRODUCTION

The retina is a sensory tissue that converts light into neural signals for vision. It consists of ten layers, from the internal limiting membrane (ILM) to the retinal pigment epithelium (RPE) [1]. An epiretinal membrane (ERM), a fibrocellular layer of glial cells, fibroblasts, or macrophages, can form on the ILM, distorting retinal layers and impairing vision [2]. ERM surgery, involving ERM removal with or without ILM peeling, improves retinal structure and function [3]. While ILM removal may reduce ERM recurrence [4], it is linked to micro-scotomas, reduced retinal sensitivity, and delayed visual recovery [5], making it a debated topic. Optical coherence tomography (OCT) is commonly used for ERM assessment, but distinguishing ILM removal from postoperative scans remains challenging and often requires surgical notes.

This work explores the use of artificial intelligence (AI) to analyze postoperative OCT scans and determine ILM removal, offering a tool for clinical follow-up and insight into structural differences between surgical techniques. To the best of our knowledge, this is the first study to develop a convolutional neural network (CNN) model for determining ERM surgery types using postoperative OCT scans. ResNet18 [6] was used as the baseline model with center OCT scans as input. To enhance accuracy, we introduced two preprocessing steps: energy-crop and wavelet denoising. Building on prior work [7], [8], we incorporated two tunable wavelet units: OrthoLatt-UwU, enforcing orthogonality and perfect reconstruction through an orthogonal lattice structure, and PR-Relax-UwU, relaxing the perfect reconstruction constraint. These units replaced traditional stride-convolution, downsampling, and pooling functions in ResNet18. Unlike max-pooling [9] and

average-pooling [10], which discard fine-grained details, these units preserve both low-pass and high-pass features, improving the model's ability to distinguish between ERM-only and ERM-ILM removal surgeries. AI classifications were compared with those of a trained ophthalmologist, who achieved 50% accuracy. The best ResNet18 model using original OCT scans reached 66%, while preprocessing improved accuracy to 72%. Further integration of OrthoLatt-UwU and PR-Relax-UwU increased accuracy to 76% and 78%, respectively. These results demonstrate that CNN-based AI models outperform human graders and highlight the effectiveness of preprocessing and tunable wavelet unit integration in postoperative OCT analysis. This is also the first study to employ tunable wavelets in AI models for medical imaging applications.

## II. LITERATURE REVIEW

Many studies indicate that ILM removal can alter the inner retinal layers, potentially affecting the integrity and function of remaining retinal cells [11]. Macular surgical techniques continue to advance toward minimizing tissue trauma, ensuring stable anatomical and visual outcomes, and preventing recurrence [12], [13]. With AI increasingly integrated into medical practice, its potential to refine surgical techniques is gaining recognition [14]. This capability is particularly useful for detecting subtle OCT changes that human observers might overlook and for patients transitioning to a new physician without access to prior surgical notes. Current research on tissue changes after ERM peeling relies primarily on histological analysis of ex vivo samples, limiting real-time retinal assessment during surgery [15], [16]. While histopathology provides valuable post-surgical insights, it is impractical for routine use. In contrast, OCT is widely accessible and commonly used both preoperatively and for follow-up, making it an ideal tool for retinal evaluation [17], [18]. Hence, AI-assisted analysis of postoperative OCT scans could help determine whether ERM alone or both ERM and ILM were removed, offering a valuable tool for retina surgeons. However, to the best of our knowledge, no prior studies have employed AI to detect structural differences between ERM-only and ERM-ILM removal using postoperative OCT scans.

## III. METHOD

An OCT scan processing pipeline was developed to classify ERM-only and ERM-ILM removal surgeries, as shown in Fig. 1. The pipeline consists of preprocessing and a CNN-based classification model. Preprocessing includes energy-crop, which removes non-essential regions using vertical pixel energy, and wavelet denoising, which filters high-frequency noise while preserving details. ResNet18 serves as the base CNN model, enhanced with OrthoLatt-UwU and PR-Relax-UwU units to retain fine details and high-frequency features, reducing information loss from max-pooling.

### A. Preprocessing Procedure

1) *Energy-Crop*: The energy-crop method is designed to remove image areas with little meaningful information, as

essential structural details are typically concentrated near the center of an OCT scan. This process uses an adaptive approach to select the relevant area by applying a threshold based on the pixel energy for each vertical coordinate. The pixel energy  $E_y$  at a vertical coordinate  $y$  is defined as follows:

$$E_y = \sum_x |p_{x,y}|, \quad (1)$$

where  $p_{x,y}$  is the pixel value at coordinate  $(x, y)$  in which  $x$  and  $y$  represent the horizontal and vertical coordinate indices, respectively. The thresholding value for cropping the image is defined as  $Threshold = mean - a \times StandardDeviation$ , where  $mean$  and  $StandardDeviation$  represent the average and standard deviation of all  $E_y$ , with the controlling parameter  $a$ . A higher  $a$ -value results in a smaller cropped area. Multiple  $a$ -values were tested in order to find the best set-up. Using this threshold, the first and last vertical coordinates where  $E_y$  exceeds  $Threshold$  are identified to define the cropping boundaries. This process is illustrated in Fig. 2.

2) *Wavelet-Denoising*: The wavelet denoising process is proposed to reduce high-frequency noise while preserving detailed scan features. The process include discrete wavelet transform (DWT) decomposition of the input image into a low-pass component  $LL$  and three high-pass components  $LH$ ,  $HL$ , and  $HH$ . The denoised image  $I_{denoised}$  is defined as the sum of  $LL$ ,  $LH$ , and  $HL$ :

$$I_{denoised} = LL + LH + HL. \quad (2)$$

The process emphasizes the  $HL$  and  $LH$  components, as they encapsulate the majority of fine structural details present in the OCT scan. In addition, the  $HH$  component is discarded, as it contains high-frequency information in both vertical and horizontal orientations, which may represent noise. The process is illustrated in Fig. 3.

### B. Tunable Wavelet Unit based Models for Surgery Classification

Tunable wavelet units are integrated into CNN models to enhance classification performance, addressing the loss of fine-grained features caused by max-pooling and average pooling in conventional CNNs. These units employ DWT decomposition with tunable coefficients, enabling the network to retain high-frequency features while optimizing filter coefficients for the classification task.

1) *Orthogonal Lattice Structure Tunable Wavelet Unit*: In OrthoLatt-UwU implementation [7], the tunable wavelet unit with the low-pass filter  $\mathbf{H}_0$  and high-pass filter  $\mathbf{H}_1$  is constructed with orthogonal lattice structure. The structure can be expressed as follows:

$$\begin{aligned} \begin{bmatrix} \mathbf{H}_0(z) \\ \mathbf{H}_1(z) \end{bmatrix} &= \begin{bmatrix} \mathbf{H}_0(z) \\ -z^{-L}\mathbf{H}_0(-z^{-1}) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{R}_K \Lambda(z^2) \cdots \mathbf{R}_1 \Lambda(z^2) \mathbf{R}_0 \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}, \end{aligned} \quad (3)$$

where  $\mathbf{R}_k$  is a rotation matrix used to construct the filter, with  $k = 0, \dots, K$ . The delay matrices within the filter

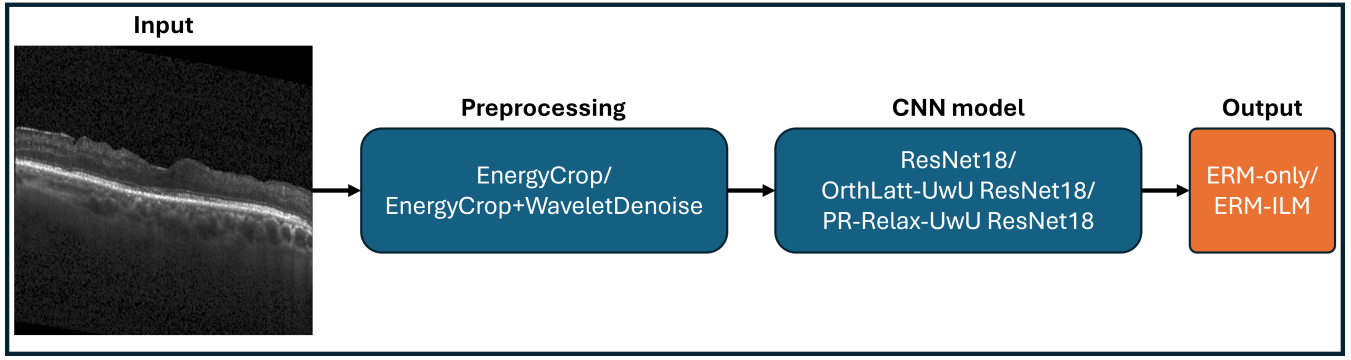


Fig. 1. Visualization of the ERM-only/ERM-ILM removal surgery classification pipeline based on OCT center scan input.

are denoted as  $\Lambda(z^2)$ . Additionally,  $N$  represents the order of the filter, defined as  $N = 2K + 1$ . In addition, rotation matrices are inherently orthogonal. Thus,  $\mathbf{R}_k$  and  $\Lambda(z^2)$  can be mathematically expressed as follows:

$$\mathbf{R}_k = \begin{bmatrix} \cos(\theta_k) & \sin(\theta_k) \\ -\sin(\theta_k) & \cos(\theta_k) \end{bmatrix} = \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix};$$

$$\Lambda(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}. \quad (5)$$

In Eq. (III-B1),  $\theta_k$  represents a rotation angle that determines the coefficients of the wavelet filter bank, with  $k = 0, \dots, K$ . These rotation angles, used in the rotation matrices—where either their rows or columns are orthonormal—are also referred to as lattice coefficients, defining the filter coefficients within the filter bank. This orthonormality ensures that the resulting filters, obtained through the multiplication of rotation and delay matrices, preserve orthogonality and hence, perfect reconstruction. These lattice coefficients  $\theta_k$  are also the tunable parameters, making the wavelet unit tunable in CNN.

2) *Tunable Wavelet Unit with Perfect Reconstruction Relaxation*: The tunable wavelet in the PR-Relax-UwU implementation [8] has  $\mathbf{H}_0$  and  $\mathbf{H}_1$  as the low-pass and high-pass filters, respectively. Since the perfect reconstruction (PR) constraint is relaxed, the filter coefficients satisfy the alias cancellation

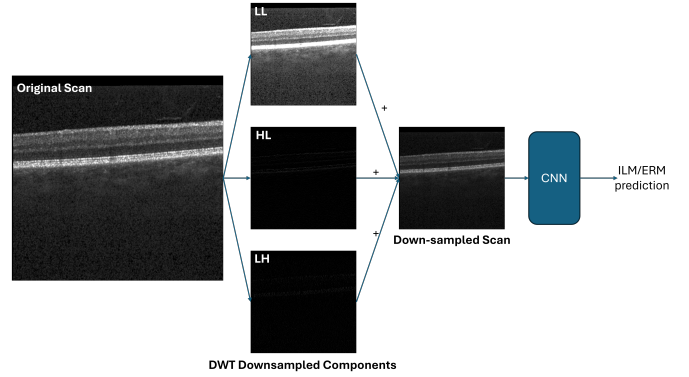


Fig. 3. Visualization of the Wavelet Denoising process. The input image is decomposed using Discrete Wavelet Transform (DWT) into an approximation component  $LL$ , a vertical detail component  $LH$ , a horizontal detail component  $HL$ , and a diagonal detail component  $HH$ . The denoised output is obtained by summing  $LL$ ,  $HL$ , and  $LH$ .

condition along with a relaxed Half-band condition. Under the satisfied alias cancellation condition, the relationship between  $\mathbf{H}_0$  and  $\mathbf{H}_1$  can be expressed as follows [19]:

$$\mathbf{h}_1(n) = (-1)^n \mathbf{h}_0(N - 1 - n), \quad (6)$$

where  $\mathbf{h}_0$ , and  $\mathbf{h}_1$  are filter coefficients of  $\mathbf{H}_0$ , and  $\mathbf{H}_1$ , respectively. In addition, to fulfill the PR constraint, the filter coefficients must satisfy the Half-band condition, which is given as follows [19]:

$$\mathbf{P}(z) + \mathbf{P}(-z) = 2, \quad (7)$$

where  $\mathbf{P}(z) = \mathbf{H}_0(z)\mathbf{H}_0(z^{-1})$ . From (7), the condition for the filter coefficients can be expressed as follows:

$$\begin{cases} \sum_{n=0}^{N-1} h(n)^2 = 1 \\ \sum_{0 < n, n+2l < N-1} h(n)h(n+2l) = 0, \text{ for } 0 < l \leq N/2. \end{cases} \quad (8)$$

From (8), the loss function for the PR constraint based on the Half-band condition can be mathematically expressed as

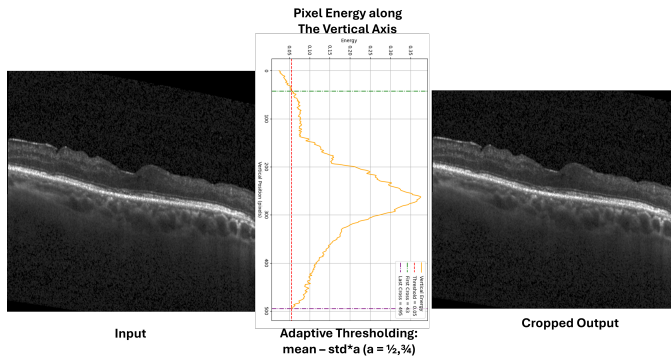


Fig. 2. Visualization of the Energy-Crop procedure. The image cropping process is performed using a threshold value defined as  $Threshold = mean - a \times StandardDeviation$ , where  $a$  is the control parameter.

TABLE I

SURGERY CLASSIFICATION PERFORMANCE BASED ON ACCURACY WITH DIFFERENT PREPROCESSING PROCEDURES: ENERGY-CROP (EC) (WITH  $\alpha = 0.75$  AND  $\alpha = 0.5$ ) AND WAVELET-DENOISING (DENOISE) (USING HAAR AND DB2 WAVELET FUNCTIONS). THE BASELINE PERFORMANCE REFERS TO RESNET18 WITHOUT PREPROCESSING. THE PERFORMANCE OF THE TRAINED HUMAN GRADER IS 50%.

Preprocessing steps	Validation	Test
<b>Baseline</b>	60.29( $\pm 2.08$ )%	58.8( $\pm 7.86$ )%
<b>EC-0.75</b>	64.41( $\pm 3.77$ )%	62( $\pm 6.81$ )%
<b>EC-0.75 + Denoise-Haar</b>	67.16( $\pm 1.39$ )%	61.33( $\pm 4.11$ )%
<b>EC-0.75 + Denoise-DB2</b>	64.22( $\pm 2.5$ )%	64( $\pm 2.83$ )%
<b>EC-0.5</b>	61.27( $\pm 3.02$ )%	62.67( $\pm 4.11$ )%
<b>EC-0.5 + Denoise-Haar</b>	61.76( $\pm 1.20$ )%	62.67( $\pm 1.89$ )%
<b>EC-0.5 + Denoise-DB2</b>	62.65( $\pm 2.00$ )%	<b>64.67</b> ( $\pm 5.25$ )%

TABLE II

BEST SURGERY CLASSIFICATION PERFORMANCES BASED ON ACCURACY USING DIFFERENT PROPOSED METHODS: ENERGY-CROP (EC) AND WAVELET-DENOISING (DENOISE) PREPROCESSING PROCEDURES, ORTHLATT-UWU, AND PR-RELAX-UWU. THE BASELINE PERFORMANCE REFERS TO RESNET18 WITHOUT PREPROCESSING. THE PERFORMANCE OF THE TRAINED HUMAN GRADER IS 50%.

Methods	Validation	Test	Best on Test
<b>Baseline</b>	60.29( $\pm 2.08$ )%	58.8( $\pm 7.86$ )%	66%
<b>EC-0.5</b>	61.27( $\pm 3.02$ )%	62.67( $\pm 4.11$ )%	68%
<b>EC-0.5&amp;Denoise-DB2</b>	62.65( $\pm 2.00$ )%	64.67( $\pm 5.25$ )%	72%
<b>OrthLatt-UwU-8Tap EC-0.5&amp;Denoise-DB2</b>	62.26( $\pm 4.86$ )%	<b>69.33</b> ( $\pm 4.99$ )%	<b>76%</b>
<b>PR-Relax-UwU-8Tap EC-0.5&amp;Denoise-DB2</b>	64.71( $\pm 2.40$ )%	<b>70.67</b> ( $\pm 5.25$ )%	<b>78%</b>

follows:

$$\mathbf{L}_{PR} = |1 - \sum_{n=0}^{N-1} h(n)|^2 + \sum_{l=1}^{N/2} \left( \sum_{0 < n, n+2l < N-1}^{N-1} h(n)h(n+2l) \right)^2. \quad (9)$$

From (6) and (9), the PR constraint is implemented to train the filter bank analysis. The relaxation of the PR constraint is achieved by multiplying  $\mathbf{L}_{PR}$  with a factor  $\alpha$ . Cross-Entropy loss  $\mathbf{L}_{CE}$  is used. A higher  $\alpha$  strengthens the Half-band constraint, while relaxing it allows for greater coefficient fine-tuning. The total loss function  $\mathbf{L}$  is expressed as follows:

$$\mathbf{L} = \mathbf{L}_{CE} + \alpha \mathbf{L}_{PR}. \quad (10)$$

For both tunable wavelet unit cases, given the low-pass and high-pass filters  $H_0(z)$  and  $H_1(z)$ , the corresponding high-pass and low-pass filter matrices are computed and denoted as  $\mathbf{H}$  and  $\mathbf{L}$ . These matrices are used to extract the approximation component  $\mathbf{X}_{ll}$  and the detail components  $\mathbf{X}_{lh}$ ,  $\mathbf{X}_{hl}$ , and  $\mathbf{X}_{hh}$  from an image or feature map  $\mathbf{X}$ . The computation of  $\mathbf{L}$  can be mathematically described as follows:

$$\mathbf{L} = \mathbf{D}\hat{\mathbf{H}}, \quad (11)$$

where  $\mathbf{D}$  is the downsampling matrix with a factor of 2, and  $\hat{\mathbf{H}}$  is a Toeplitz matrix with filter coefficients  $h_0, h_1, \dots, h_{2N-1}$  of  $H_0(z)$ . The matrix  $\mathbf{H}$  has a similar structure to  $\mathbf{L}$  but uses the filter coefficients of  $H_0(z^{-1})$ . Using  $\mathbf{H}$  and  $\mathbf{L}$ , the components  $\mathbf{X}_{ll}$ ,  $\mathbf{X}_{lh}$ ,  $\mathbf{X}_{hl}$ , and  $\mathbf{X}_{hh}$  are computed as follows:

$$\begin{aligned} \mathbf{X}_{ll} &= \mathbf{L}\mathbf{X}\mathbf{L}^T, & \mathbf{X}_{lh} &= \mathbf{H}\mathbf{X}\mathbf{L}^T, \\ \mathbf{X}_{hl} &= \mathbf{L}\mathbf{X}\mathbf{H}^T, & \mathbf{X}_{hh} &= \mathbf{H}\mathbf{X}\mathbf{H}^T. \end{aligned} \quad (12)$$

In this work, we use a one-layer Fully Convolutional Network (FCN) to combine features from the sub-sample low-pass and high-pass components extracted via the Discrete Wavelet Transform (DWT). The combined features can be mathematically expressed as follows:

$$X_p = F'(\text{ReLU}(\mathbf{X}_{ll}), \text{ReLU}(\mathbf{X}_{lh}), \text{ReLU}(\mathbf{X}_{hl}), \text{ReLU}(\mathbf{X}_{hh})), \quad (13)$$

where  $F'$  is the one-layer FCN with tunable weights. The tunable wavelet units are integrated into the CNN architectures. For the downsampling and pooling layers, we use the decomposed components as inputs for a one-layer Fully Convolutional Network (FCN). In addition, we substitute the two-stride convolution with a non-stride convolution block followed by the tunable wavelet units.

#### IV. EXPERIMENTS AND RESULTS

This study adhered to the Declaration of Helsinki and HIPAA regulations. Written informed consent was obtained for surgery, and all patient data were anonymized. OCT scans were reviewed within 12 weeks post-surgery for quality and completeness, with surgical notes determining ILM peeling status. Anonymized SD-OCT scans were imported as video files for human grading and RAW data for AI analysis. The dataset comprises of 340 training-validation samples and 50 test samples, maintaining a balanced 50/50 ERM-only and

ERM-ILM distribution. A human grader was trained and tested on a shuffled set (25 eyes per group) without access to surgical notes. The training-validation set was randomly shuffled three times (3-fold experiments) and split 80/20 while preserving the 50/50 balance for the surgery classes. Models were trained on different training-validation splits and evaluated on the same test set, with the trained ophthalmologist achieving 50% accuracy. All models were trained using the Adam optimizer with a 0.001 learning rate for 200 epochs and a batch size of 128. The preprocessing methods, energy-crop (EC) and wavelet-denoising (Denoise), were evaluated against the baseline ResNet18 model trained on original data. EC experiments tested  $\alpha$ -values of 0.5 and 0.75, with the best setup applied to Denoise experiments using Haar and DB2 wavelets. Results are reported in Table I.

Based on test set performance in Table I, the best setup (EC-0.5 with Denoise-DB2) was selected for tunable wavelet unit integration experiments with OrthLatt-UwU and PR-Relax-UwU. The tunable wavelet units having filters with 2-Tap (Haar), 4-Tap (DB2), 6-Tap (DB3), and 8-Tap (DB4) were tested, with only the best-performing 8-Tap models reported in Table II. Table II presents the mean accuracies and best test set performance for each method. From Table I and Table II, AI models consistently outperformed trained human graders in distinguishing between ERM-alone and ERM-ILM removal surgeries. The combination of preprocessing and wavelet-based enhancements significantly improved model performance, demonstrating AI's potential in refining postoperative assessment and supporting clinical decision-making. Additionally, relaxing the perfect reconstruction constraint in tunable wavelet unit integration yielded the best classification results.

## V. CONCLUSION

This study is the first to develop convolutional neural network (CNN) models for classifying ERM surgery types using postoperative OCT scans. Using ResNet18 as a baseline, we introduced preprocessing techniques, including energy-crop and wavelet downsampling, to improve accuracy. Additionally, tunable wavelet units—OrthoLatt-UwU and PR-Relax-UwU—were integrated to enhance feature extraction and classification. Our results show that AI models consistently outperform trained human graders in distinguishing between ERM-alone and ERM-ILM removal surgeries. The combination of preprocessing and wavelet-based enhancements significantly improved model performance, highlighting AI's potential in refining postoperative assessment and clinical decision-making. Enhancing OCT analysis with AI offers a non-invasive, cost-effective approach for assessing retinal structure, optimizing surgical techniques, understanding disease pathology, and classifying recurrence. In future work, the current models can be extended to process OCT volumes as inputs instead of single OCT scans, potentially enhancing performance. Additionally, augmentation and diffusion techniques can be employed to increase training data, allowing

deeper networks to be trained without overfitting, leading to more stable models with reduced performance variability.

## REFERENCES

- [1] N. Mahabadi and Y. A. Khalili, "Neuroanatomy, retina," Aug 2023.
- [2] V. M. Kanukollu and P. Agarwal, "Epiretinal membrane," Jul 2023.
- [3] Y. Matsuoka, M. Tanito, Y. Takai, Y. Koyama, S. Nonoyama, and A. Ohira, "Visual function and vision-related quality of life after vitrectomy for epiretinal membranes: A 12-month follow-up study," *Investigative Ophthalmology & Visual Science*, vol. 53, p. 3054, May 2012.
- [4] N. Rayess, D. Vail, and P. Mruthyunjaya, "Rates of reoperation in 10 114 patients with epiretinal membranes treated by vitrectomy with or without inner limiting membrane peeling," *Ophthalmology Retina*, vol. 5, no. 7, pp. 664–669, 2021.
- [5] J.-B. Ducloyer, Y. Eude, C. Volteau, O. Lebreton, A. Bonissent, P. Fossum, R. Tadayoni, C. P. Creuzot-Garcher, Y. Le Mer, J. Perol, J. Fortin, A. Jobert, F. Billaud, C. Ivan, A. Poinas, and M. Weber, "Pros and cons of internal limiting membrane peeling during epiretinal membrane surgery: a randomised clinical trial with micropertometry (peeling)," *British Journal of Ophthalmology*, vol. 109, no. 1, pp. 119–125, 2025.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.
- [7] A. D. Le, S. Jin, Y.-S. Bae, and T. Q. Nguyen, "A lattice-structure-based trainable orthogonal wavelet unit for image classification," *IEEE Access*, vol. 12, pp. 88715–88727, 2024.
- [8] A. D. Le, S. Jin, Y. S. Bae, and T. Nguyen, "A novel learnable orthogonal wavelet unit neural network with perfection reconstruction constraint relaxation for image classification," in *2023 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pp. 1–5, 2023.
- [9] J. Weng, N. Ahuja, and T. Huang, "Cresceptron: a self-organizing neural network which grows adaptively," in *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, vol. 1, pp. 576–581 vol.1, 1992.
- [10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [11] R. Gelman, W. Stevenson, C. Prospero Ponce, D. Agarwal, and J. B. Christoforidis, "Retinal damage induced by internal limiting membrane removal," *Journal of Ophthalmology*, vol. 2015, no. 1, p. 939748, 2015.
- [12] M. Ferrara, A. Rivera-Real, R. J. Hillier, M. Habib, M. R. Kadhim, M. T. Sandinha, K. Curran, A. Muldrew, and David, "A randomised controlled trial evaluating internal limiting membrane peeling forceps in macular hole surgery," *Graefes Archive for Clinical and Experimental Ophthalmology*, vol. 261, p. 1553–1562, Dec 2022.
- [13] R. Gaber, Q. S. You, I. K. Muftuoglu, M. Alam, F. F. Tsai, N. Mendoza, and W. R. Freeman, "Characteristics of epiretinal membrane remnant edge by optical coherence tomography after pars plana vitrectomy," *Retina*, vol. 37, p. 2078–2083, Jan 2017.
- [14] S. Narayanan, R. Ramakrishnan, E. Durairaj, A. Das, S. Narayanan, R. Ramakrishnan, E. Durairaj, and A. Das, "Artificial intelligence revolutionizing the field of medical education," *Cureus*, vol. 15, Nov 2023.
- [15] Özlem Beyazyildiz, M. H. Tırış, E. R. Hekimoğlu, E. Beyazyildiz, F. Kaymaz, P. Yilmazbaş, and F. Öztürk, "Histopathological analysis of internal limiting membrane surgically peeled from eyes with epiretinal membrane," *Current Eye Research*, vol. 41, no. 2, pp. 258–265, 2016. PMID: 25848907.
- [16] S. K. Gibran, B. Flemming, T. Stappler, I. Pearce, C. Groenewald, H. Heimann, P. Hiscott, and D. Wong, "Peel and peel again," *British Journal of Ophthalmology*, vol. 92, p. 373–377, Nov 2007.
- [17] M. Zeppieri, S. Marsili, E. S. Enaholo, A. O. Shuaibu, N. Uwagboe, C. Salati, L. Spadea, and M. Musa, "Optical coherence tomography (oct): A brief look at the uses and technological evolution of ophthalmology," *Medicina*, vol. 59, p. 2114, Dec 2023.
- [18] E. S. Enaholo, M. J. Musa, and M. Zeppieri, "Optical coherence tomography," Oct 2024.
- [19] G. Strang and T. Q. Nguyen, "Wavelets and filter banks," 1996.