MACHINE LEARNING FOR RETINAL HEALTH CLASSIFICATION OF OPTICAL COHERENCE TOMOGRAPHY IMAGES

BY

METHAWEE RATANAPUNPERM

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE (ENGINEERING AND TECHNOLOGY)
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2025

THAMMASAT UNIVERSITY
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY

THESIS

BY

METHAWEE RATANAPUNPERM

ENTITLED

MACHINE LEARNING FOR RETINAL HEALTH CLASSIFICATION
OF OPTICAL COHERENCE TOMOGRAPHY IMAGES

was approved as partial fulfillment of the requirements for
the degree of Master of Science (Engineering and Technology)

on December 9, 2025

Chairperson _____
(Duangrat Gansawat, Ph.D.)

Member and Advisor _____
(Associate Professor Pakinee Aimmanee, Ph.D.)

Member _____
(Assistant Professor Somrudee Deepaisarn, Ph.D.)

Director _____
(Associate Professor Kriengsak Panuwatwanich, Ph.D.)

Thesis Title	MACHINE LEARNING FOR RETINAL HEALTH CLASSIFICATION OF OPTICAL COHERENCE TOMOGRAPHY IMAGES
Author	Methawee Ratanapunperm
Degree	Master of Science (Engineering and Technology)
Faculty/University	Sirindhorn International Institute of Technology/ Thammasat University
Thesis Advisor	Associate Professor Pakinee Aimmanee, Ph.D.
Academic Years	2025

ABSTRACT

Optical Coherence Tomography (OCT) is an essential imaging technique for diagnosing retinal diseases. While deep learning models offer high accuracy for automated OCT classification, their "black-box" nature limits clinical trust and adoption. Conversely, traditional machine learning methods are more interpretable but often depend on features that can be difficult to extract or are sensitive to image quality. This thesis presents two complementary studies to develop an accurate and interpretable machine learning framework for retinal health classification.

The first study introduces a method based on extracting computationally simple and clinically intuitive geometric features—specifically foveal concavity, bilateral symmetry, and layer smoothness—from the Inner Limiting Membrane (ILM) and Retinal Pigment Epithelium (RPE) layers. Using a Light Gradient Boosting Machine (LGBM) classifier, this approach achieved 96% accuracy, 98% precision, and a low false negative rate of 6%, significantly outperforming a baseline method using vertical line profile features.

The second study expands this analysis by extracting a comprehensive set of 56 features derived from the thickness, area, and inter-layer relationships of seven segmented retinal layers. This layer-wise approach, while achieving a slightly lower binary classification accuracy of 93.5% on the test set, demonstrated robust performance and provided deeper clinical insights. Through interpretable models, this analysis revealed a hierarchical pattern of retinal

deterioration. Key predictive features included the area ratio of specific layers to the total retina and thickness ratios between adjacent layers, offering a more granular understanding of disease mechanisms.

Together, these studies demonstrate that analyzing retinal layer morphology, from simple geometries to complex inter-layer dynamics, provides a powerful and transparent framework for automated disease detection. The proposed methods support ophthalmologists by improving diagnostic accuracy and efficiency, contributing to the early detection and prevention of vision loss. The initial geometric approach is ideal for rapid pre-screening, while the layer-wise analysis enhances diagnostic insight by uncovering specific, clinically meaningful patterns of pathology.

Keywords: Unhealthy OCT Image Detection, Machine Learning, Retinal Layer Geometry, Retinal Layer Relationships, Interpretability, Feature Engineering.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to Sirindhorn International Institute of Technology for granting me the Faculty's Quota Scholarship, which provided crucial financial support and enabled me to pursue my graduate studies. I am also thankful for the necessary resources, infrastructure, and academic environment that facilitated this research. I am truly grateful for the opportunity and the trust placed in me.

I would also like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Pakinee Aimmanee, for her insightful feedback, experienced guidance, and heart-warming support throughout this research journey. Her expertise and supportiveness have been greatly significant in shaping this dissertation.

I am sincerely thankful to Asst. Prof. Dr. Somrudee Deepaisarn for her detailed and constructive suggestions and support during the course of this work. Her suggestions helped improve my research strategy and greatly enhanced the quality of this thesis.

I wish to extend my genuine thanks to Dr. Duangrat Gansawat for her invaluable guidance and insightful advice. Her expertise, thoughtful feedback, and unwavering encouragement were instrumental in the development and refinement of this work.

A special thanks to my colleague, Mr. Thanaphat Khemniwat, for his camaraderie, stimulating discussions, and moral support during the challenging phases of this work. His contributions made this study possible.

Finally, I owe my deepest gratitude to my family and friends for their unwavering love, encouragement, and sacrifices. Their belief in me kept me motivated throughout this academic journey.

This dissertation would not have been possible without the collective support of all these individuals and institutions.

Methawee Ratanapunperm

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

Symbols	Terms
$\%$	percentage
t	threshold
ω	class probability
σ	standard deviation
a, b, c	quadratic coefficients
y	vertical coordinate
x	horizontal coordinate
n	number of samples
TP	True Positives
TN	True Negatives
FP	False Positives
FN	False Negatives

Abbreviations	Terms
AdaBoost	Adaptive Boosting
AMD	Age-Related Macular Degeneration
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolutional Neural Network
CNV	Choroidal Neovascularization
CSC	Central Serous Chorioretinopathy
CSR	Central Serous Retinopathy
DL	Deep Learning
DME	Diabetic Macular Edema
DR	Diabetic Retinopathy
ILM	Inner Limiting Membrane
kNN	k-Nearest Neighbors
LG	Machine Learning Using Layer Geometry Features
LGBM	Light Gradient Boosting Machine
LP	Machine Learning Using Line Profile Features
LSR	Machine Learning Using Layer Structural and Relational Features
ME	Macular Edema
ML	Machine Learning
OCT	Optical Coherence Tomography
OSL	Outer Segment Layer
RE	Retinal Edema
RFC	Random Forest Classifier
RMSE	Root Mean Square Error
RPE	Retinal Pigment Epithelium
SVM	Support Vector Machine

CHAPTER 1
INTRODUCTION

1.1 Background

OCT is a non-invasive, high-resolution imaging technique that captures cross-sectional and volumetric images of biological tissues, particularly the retina. By utilizing low-coherence interferometry, OCT provides micrometer-scale resolution, enabling clinicians to visualize retinal layers, detect abnormalities, and monitor disease progression in real time. Due to its speed, safety, and diagnostic precision, OCT has become a gold standard in ophthalmology for managing a variety of retinal diseases, such as age-related macular degeneration (AMD), diabetic retinopathy (DR), and macular edema. With the ability to show retinal layers, OCT imaging is essential for early diagnosis and therapy planning.

A normal OCT scan, extending from the vitreous interface to the choroid, reveals clearly delineated retinal layers ranging from the ILM down to the RPE. In such scans, the retinal architecture remains intact, with no signs of fluid buildup, cystic formations, or layer disorganization. The foveal depression is distinctly observable in healthy scans. As shown in Figure 1.1(A), the ILM, RPE boundary, and foveal contour are clearly marked.

Conversely, OCT scans of pathological conditions display notable deviations from this structure. For instance, choroidal neovascularization (CNV) is characterized by hyperreflective, irregular formations beneath the RPE, often accompanied by fluid accumulation or cystic changes. Diabetic macular edema (DME) typically presents with intraretinal cysts, diffuse retinal thickening, and sometimes subretinal fluid, leading to a distorted retinal profile. OCT images showing drusen often reveal elevated, hyperreflective lesions beneath the RPE with disruptions in the outer retinal layers. Figure 1.1(B) provides visual examples of these abnormal retinal features.

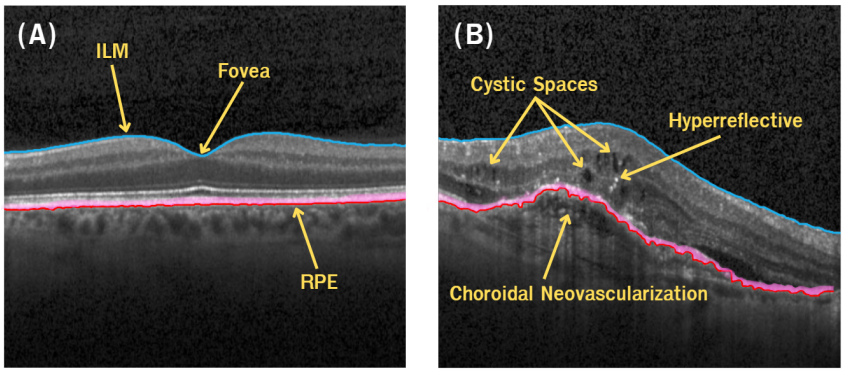

Figure 1.1 (A) Example of a normal OCT image, and (B) example of an abnormal OCT image. Pink stripe indicates the RPE layer. The ILM layer and the bottom edge of the RPE layer are indicated with blue and red lines, respectively.

OCT image segmentation further enhances the understanding of retinal structure by delineating the seven internal retinal layers. When these intraretinal layers are segmented, the smooth, regular patterns of a normal retina become evident, reflecting intact retinal architecture and consistent layer thickness. In contrast, abnormal OCT scans exhibit severely disrupted and irregular layer boundaries, with distortions and discontinuities that correspond to pathological changes. This detailed segmentation allows for precise visualization of how disease processes affect each retinal layer, aiding in diagnosis and monitoring. Figure 1.2 illustrates this contrast, showing segmented layers of a normal OCT image in (C) with clear, continuous layers, and an abnormal OCT image in (D) where the layered structure is markedly disturbed.

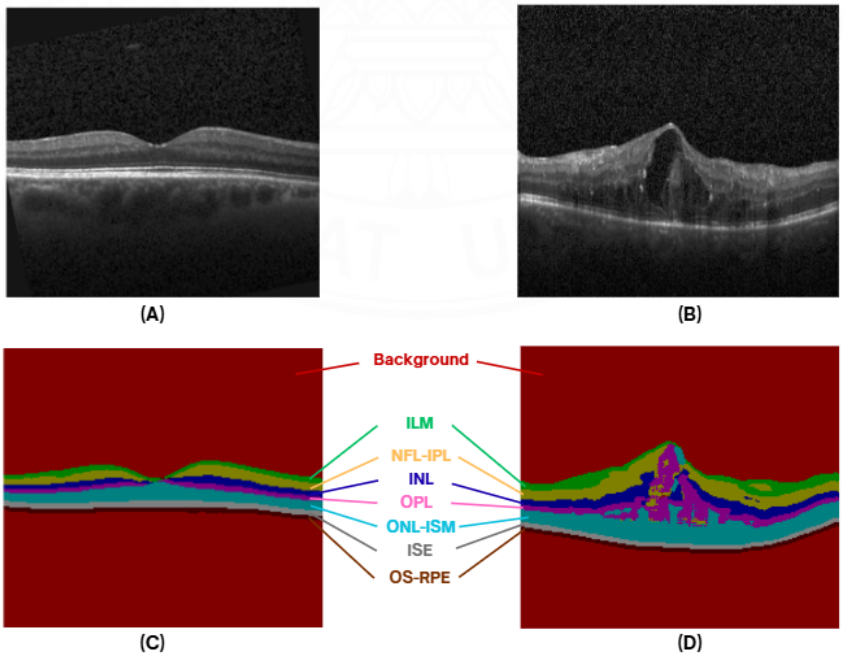

Figure 1.2 (A) Example of a normal OCT image, and (B) example of an abnormal OCT image. (C) Segmented layers of a normal OCT image, and (D) segmented layers of an abnormal OCT image.

1.2 Problem Statement

OCT is a vital diagnostic tool in ophthalmology, enabling high-resolution cross-sectional imaging of retinal structures. However, manual interpretation of OCT scans by medical professionals is time-consuming, subjective, and prone to inter-observer variability. To address these challenges, numerous ML and DL approaches have been proposed for the automated classification of retinal OCT images into healthy and pathological categories.

Although DL-based methods such as Convolutional Neural Networks (CNNs), ResNet, and VGG have demonstrated high accuracy in OCT classification, they suffer from several critical limitations. These models operate as complex, opaque systems—commonly referred to as having a ”black-box” nature—making it difficult for clinicians to understand how decisions are made. This lack of interpretability reduces trust in automated systems, particularly in medical diagnostics where explainability is essential. Furthermore, DL models require large annotated datasets for training, which are often limited in medical imaging due to privacy concerns and the need for expert labeling. Additionally, many DL architectures demand substantial computational resources, making them challenging to deploy in resource-constrained clinical settings.

In contrast, traditional feature-based ML models, such as Support Vector Machines (SVM) and Random Forests, rely on handcrafted features based on texture, geometry, and intensity descriptors. These models are generally more interpretable with feature importance and more computationally efficient. However, many existing ML approaches still face significant challenges, including limited generalizability across diverse OCT datasets, difficulty in selecting optimal features to maximize classification performance, and the need for better integration of domain knowledge into model design.

1.3 Research Objective

This study aims to advance the understanding and automated analysis of retinal OCT images by integrating two complementary approaches. First, it seeks to develop an interpretable machine learning system that classifies retinal OCT images as healthy or unhealthy based on geometric features of the ILM and RPE layers, emphasizing clinical relevance and transparency. Second, it expands the scope by extracting and analyzing features derived from retinal layer thickness and area measurements, investigating the relationships and insights between individual retinal layers and overall retinal health.

The key objectives are:

- To extract and analyze interpretable geometric features of the ILM and RPE layers (e.g., symmetry, smoothness, layer thickness variations) that correlate with retinal abnormali-

ties, ensuring clinical relevance and transparency in decision-making.

- To extract quantitative features based on retinal layer thickness and area, and to explore the relationships and insights between these layers and retinal health status.
- To compare the performance of the retinal layer-wise relationship feature approach with the geometric feature approach, demonstrating improvements or complementary strengths.
- To validate the clinical utility of the integrated system by assessing its suitability for pre-screening applications, aiming to reduce ophthalmologists' workload while maintaining diagnostic reliability.
- To provide an open and reproducible framework adaptable to different OCT datasets, facilitating future research in retinal disease detection and interpretation.

1.4 Overview of the Thesis

This thesis is organized into six chapters to systematically develop and evaluate machine learning methods for retinal OCT image analysis. Following this introduction, Chapter 2 reviews previous research on OCT imaging and classification techniques. Chapter 3 details the two main methodologies used: first, the extraction of simple geometric features from the ILM and RPE layers, and second, a more comprehensive analysis based on thickness and area relationships across all seven retinal layers. Chapter 4 describes the experiments, presenting the performance results for both feature sets and comparing them against existing methods. Chapter 5 provides a unified discussion of these results, analyzing the complementary strengths of each approach. Finally, Chapter 6 concludes the thesis by summarizing the key findings.

CHAPTER 2
LITERATURE REVIEW

Recent advancements in ML and DL have significantly improved the automated detection of retinal abnormalities in OCT images. This chapter reviews key studies in the field, focusing on existing feature-based ML approaches and their limitations.

2.1 Structural Feature-Based Approaches for OCT Classification

Several studies have demonstrated that structural and geometric features derived from retinal anatomy can effectively classify OCT images with high accuracy while maintaining interpretability.

Hussain et al. (2018) demonstrated that retinal layer-based features—such as thickness, curvature, and pathology-associated volumes—offer superior discriminative capability and are more robust to noise compared to traditional intensity- or texture-based features. Texture descriptors, while widely used, tend to be highly sensitive to image noise and exhibit significant variability across different OCT devices due to inconsistencies in intensity scaling and resolution. In contrast, structural features remain relatively invariant across devices, improving generalizability. Their study employed a Random Forest Classifier (RFC), achieving an accuracy exceeding 96% in classifying normal, AMD, and DME cases. Furthermore, the Outer Segment Layer (OSL) thickness was found to be significantly reduced in early AMD cases, reinforcing the importance of fine structural measurements as early biomarkers for retinal degeneration.

Khalid et al. (2017) employed a multi-layered SVM classifier based on nine retinal features, including thickness and cyst/drusen presence, achieving an exceptional 99.92% accuracy in classifying healthy, Retinal Edema (RE), Central Serous Chorioretinopathy (CSC), and AMD cases.

Syed et al. (2016) introduced a 3D retinal surface reconstruction approach using coherent tensors, extracting features like thickness profiles and cyst space information. Their SVM-based model achieved 98.88% accuracy in detecting healthy, Central Serous Retinopathy (CSR), and Macular Edema (ME) images.

While these methods demonstrate strong performance, retinal thickness-based features have limitations:

- **Sensitivity to image quality:** Variations in scan conditions (e.g., noise, motion artifacts) can degrade layer segmentation accuracy.

- **Dependency on high-quality data:** Many existing techniques require accurate segmentation of retinal layers, which can be unreliable in the presence of severe retinal pathologies.
- **Reduced robustness in severe pathologies:** Diseased retinas often exhibit distorted layers, making thickness measurements unreliable.

2.2 Alternative Feature-Based Approaches for OCT Classification

To address these challenges, researchers have explored alternative feature extraction techniques.

Salaheldin et al. (2024) combined DL with traditional ML, using InceptionV3 for feature extraction followed by classifiers (SVM, k-Nearest Neighbors (kNN), Decision Tree). Their hybrid model attained over 97% accuracy in classifying CNV, DME, drusen, and healthy cases.

Numsonthi et al. (2025) proposed a computationally efficient method using vertical line-profile features which utilize intensity of the images for binary classification (healthy vs. unhealthy). Their approach achieved 91% accuracy with lower computational costs but was sensitive to noise and struggled with small pathological regions.

Chanpipattanachai et al. (2026), building on Numsonthi et al. (2025)'s work, has developed a machine learning framework using features from horizontal, vertical, and combined intensity line profiles. The model using combined horizontal and vertical profiles yielded the highest accuracy of 94%, an improvement of 2-3% over a prior state-of-the-art method that used only vertical profiles. Although horizontal features alone performed poorly, their integration with vertical features enhanced classification metrics by 1-2%. This combined approach offers an efficient and interpretable method for large-scale OCT screening, though misclassifications occurred due to image distortions and the features' limitations in capturing certain structural anomalies.

Despite their success, existing methods face critical challenges:

- **Interpretability vs. performance trade-off:** The more advanced feature extraction techniques often achieve high accuracy but are harder to understand, making them difficult to interpret and limiting their acceptance in clinical settings. In contrast, structural feature-based features offer greater interpretability but may have difficulty generalizing across diverse datasets.
- **Computational efficiency:** Feature extraction using DL approaches typically demand significant computational resources, while also producing hard-to-interpret features.

2.3 Research Gap and Proposed Contribution

This study addresses these gaps by:

- Leveraging geometric features of the ILM and RPE—such as symmetry and smoothness—that are computationally simple to extract and meaningful to clinicians, while still achieving strong classification performance.
- Designing a lightweight ML framework that balances diagnostic accuracy and computational efficiency, making it suitable for real-world clinical use.
- Conducting a comprehensive comparison with more complex feature-based models to demonstrate the effectiveness and practical benefits of using simpler features.

CHAPTER 3
METHODOLOGY

The methodology presented in this study follows the straightforward guidelines illustrated in Figure 3.1. This workflow provides a clear framework for the sequential steps involved.

For feature processing, we proposed two methods. The first method involves using the area between the ILM and RPE boundaries to extract features, while the second method utilizes each layer individually in addition to the first. Details of each one is described in the sections 3.1 and 3.2.

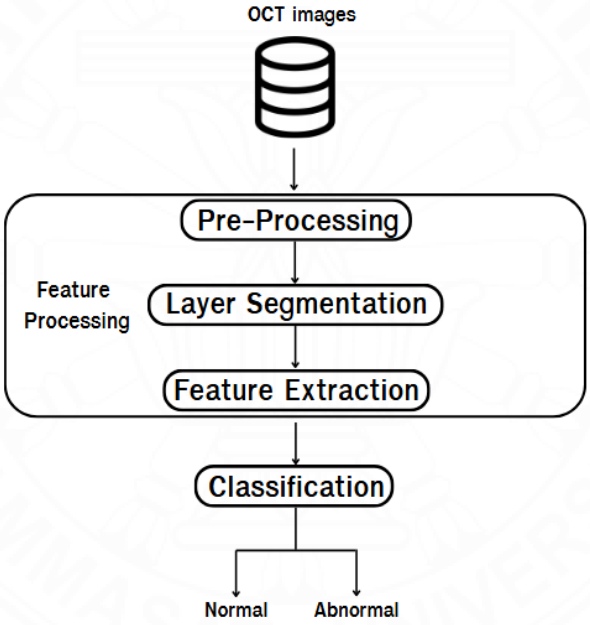

Figure 3.1 General workflow of our methodology.

3.1 Machine learning Using Geometrical Feature

For ease, we call this method that uses layer geometry as LG. Details of the method are as follows.

3.1.1 Pre-processing

The initial step in image pre-processing involves preparing the OCT scans to ensure visual clarity and structural consistency. To begin, any surrounding white margins are removed. Since this operation may result in irregular image shapes, a perspective transformation

(Szeliski, 2022) is applied to redefine a standard rectangular region. To improve image quality, a two-stage denoising process is carried out. Fast non-local means denoising (Gonzalez & Woods, 2008) is first applied to reduce random noise, followed by bilateral filtering (Gonzalez & Woods, 2008) to preserve edges and structural details while smoothing homogeneous areas.

3.1.2 Segmentation of ILM and RPE Layers

The ILM and the lower boundary of the RPE layer (Abhishek et al., 2014) are chosen as the primary anatomical landmarks due to their consistent visibility across all images, including those with pathological features.

For ILM segmentation, Otsu's thresholding method (Yousefi, 2011) is used. This method selects a threshold t that minimizes the intra-class variance between foreground and background pixels, as defined in Equation 3.1.

$$\min_{t} \quad \omega_0(t)\sigma_0^2(t) + \omega_1(t)\sigma_1^2(t), \tag{3.1}$$

where ω_0 and ω_1 represent the class probabilities, and $\sigma_0^2(t)$ and $\sigma_1^2(t)$ are the class variances. This segmentation process effectively identifies the uppermost boundary separating the dark background and the brighter retinal region, which corresponds to the ILM layer.

The RPE layer is characterized by a bright, thick profile near the bottom of the image. To segment it accurately, contrast enhancement is first applied using Contrast Limited Adaptive Histogram Equalization (CLAHE) and gamma correction (Gonzalez & Woods, 2008). These operations enhance the distinction between the RPE and adjacent layers. Otsu's thresholding is then employed again to extract the lower edge of the RPE.

Once both boundaries are segmented, the image is rotated so that the ILM and RPE layers lie approximately parallel to the horizontal axis. The required rotation angle is computed using linear regression on the RPE boundary points, and each point is rotated using the 2D rotation formula. This alignment standardizes the orientation of all images.

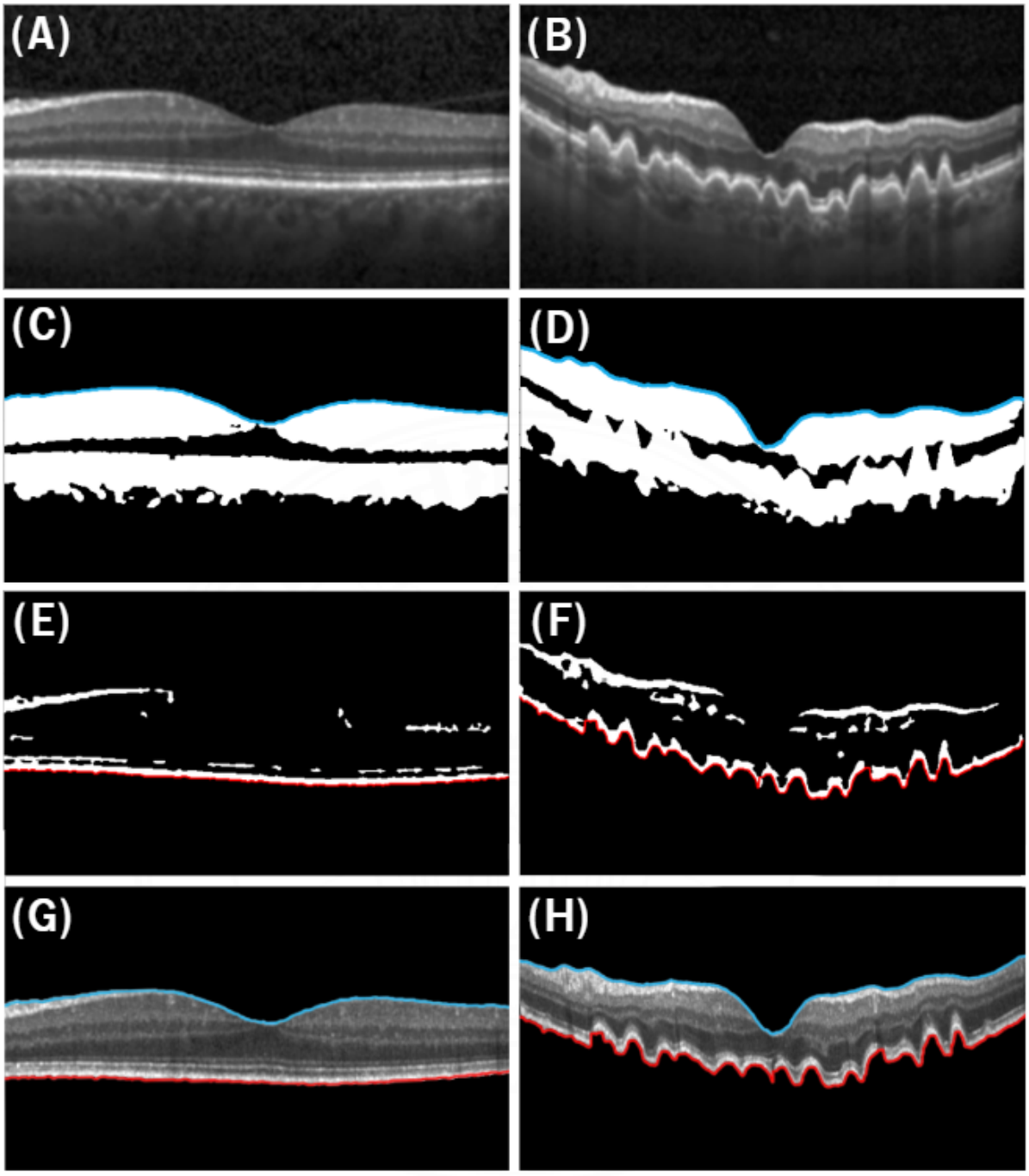

Figure 3.2 Illustration of normal (left) and abnormal (right) images after each processing step. (A)-(B): Post-denoising; (C)-(D): ILM segmentation via Otsu's method; (E)-(F): RPE segmentation; (G)-(H): Rotated layers.

3.1.3 Feature Extraction

Four features are extracted to characterize foveal geometry and layer symmetry.

3.1.3.1 Foveal Concavity

Foveal concavity is a binary feature: "concave" or "convex". The ILM boundary is smoothed using an adaptive moving average filter (Press et al., 2007), and the foveal location is detected as the nearest local extremum. Two reference points are placed symmetrically at 10% of the image width on either side. If the foveal point lies below both references, it is classified as concave; otherwise, convex.

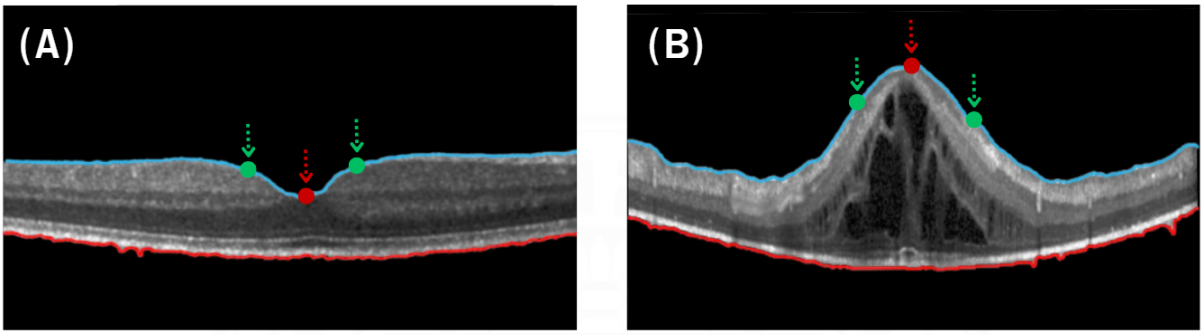

Figure 3.3 (A) Concave fovea: foveal point lies below reference points. (B) Convex fovea: foveal point lies above at least one reference point. Red dot: fovea; green dots: references.

3.1.3.2 Bilateral Symmetry

Symmetry of ILM and RPE layers is evaluated by comparing the left half with the mirrored right half across the vertical foveal line. Root Mean Square Error (RMSE) is calculated using:

$$\text{RMSE} = \sqrt{\sum_{i}^{n} \frac{(L_i - RR_i)^2}{n}}, \tag{3.2}$$

where L_i is the i^{th} y-coordinate on the left, and RR_i is the mirrored coordinate from the right side. Only overlapping regions are used when left and right segments differ in length.

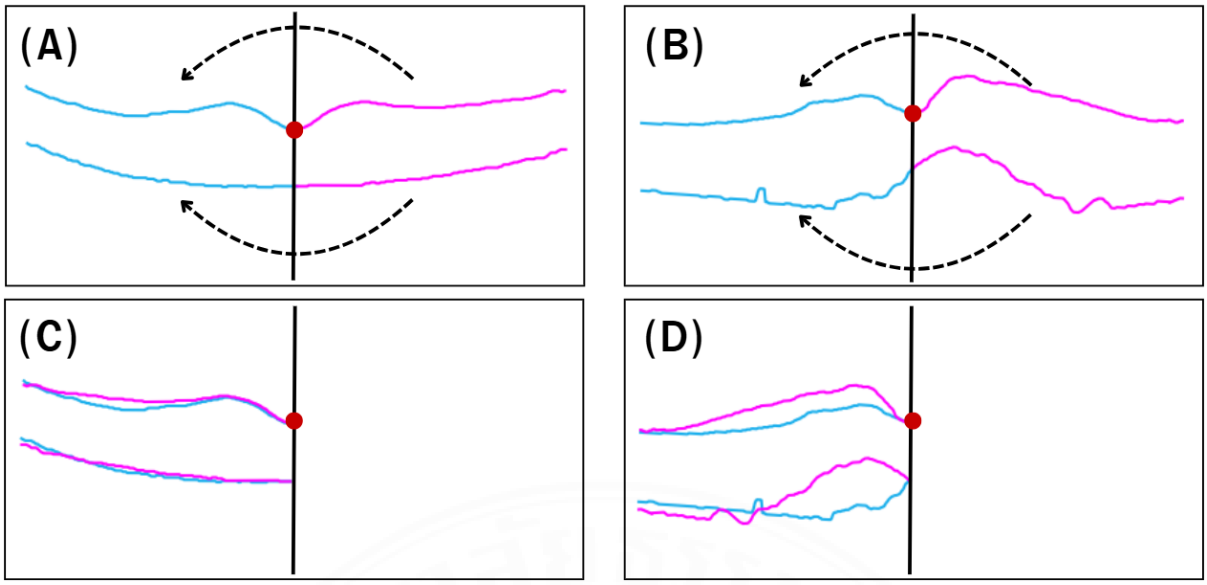

Figure 3.4 Bilateral symmetry evaluation: (A)-(B) splitting at foveal center; (C)-(D) comparison with mirrored right halves. Left: normal; Right: abnormal.

3.1.3.3 RPE Smoothness

RPE smoothness is measured by fitting the layer to a quadratic curve:

$$y = ax^2 + bx + c, \tag{3.3}$$

where a, b, and c are constants fitted to minimize the squared error. RMSE is computed between the actual RPE boundary and its quadratic fit. Lower RMSE implies higher smoothness, which typically correlates with healthy retinal conditions.

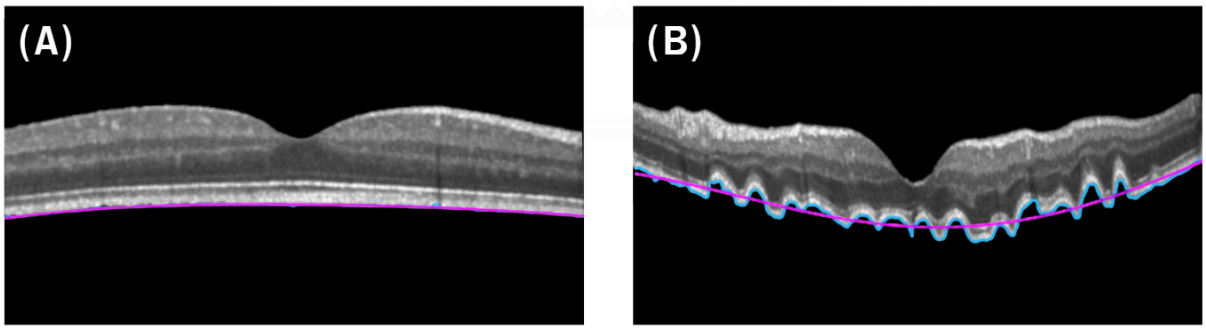

Figure 3.5 RPE (blue) and its quadratic fit (pink) for (A) normal and (B) abnormal cases. Normal profiles show tighter adherence to the fitted curve.

3.2 Machine Learning Using Layer Structural and Relational Features

We name this method that uses layer structural and relational features as LSR.

3.2.1 Pre-processing

Images of varying sizes and aspect ratios are padded to a square shape with a 1:1 aspect ratio, regardless of the original width and height, preserving structural integrity and spatial consistency without distortion. The padding value is obtained from the mode pixel value of the image to ensure that the padded pixels blend seamlessly with the background.

To correct tilted retinas, a rotation correction algorithm aligns the retina horizontally by detecting its principal axis.

Noise levels are adaptively estimated per image using the Median Absolute Deviation of high-frequency components, guiding a two-step denoising process: Gaussian filtering reduces high-frequency noise proportionally to the estimated noise, followed by fast Non-Local Means filtering to preserve structural details while further suppressing noise.

Finally, contrast enhancement is applied adaptively based on the Weber contrast metric, with iterative gamma correction adjusting pixel intensities to optimize retinal layer visibility within an empirically determined contrast range, thereby facilitating precise layer segmentation.

The preprocessing steps are shown in Figure 3.6.

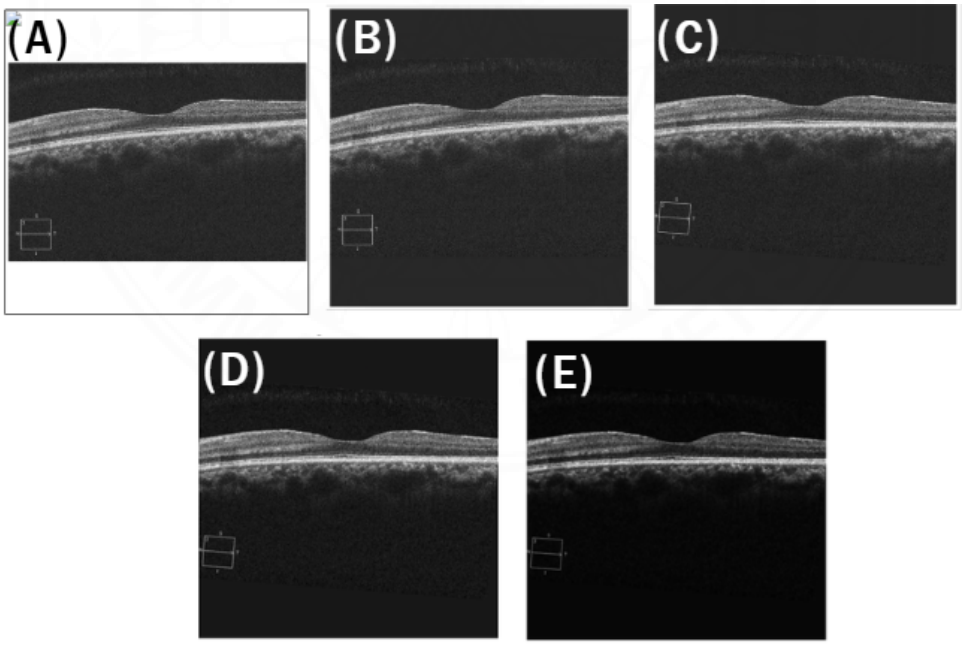

Figure 3.6 (A) Original image. (B) Image after padding with mode value. (C) Padded image after rotation. (D) Rotated image after denoising. (E) Denoised image after contrast enhancement.

3.2.2 Multi-Layer Segmentation

We utilize the Y-Net segmentation model by Farshad et al. (2022), which is designed to extract seven retinal layers along with the background and fluid regions from OCT images. In this study, segmentation is focused exclusively on the seven retinal layers and the background. The Y-Net segmentation pipeline is illustrated in Figure 3.7.

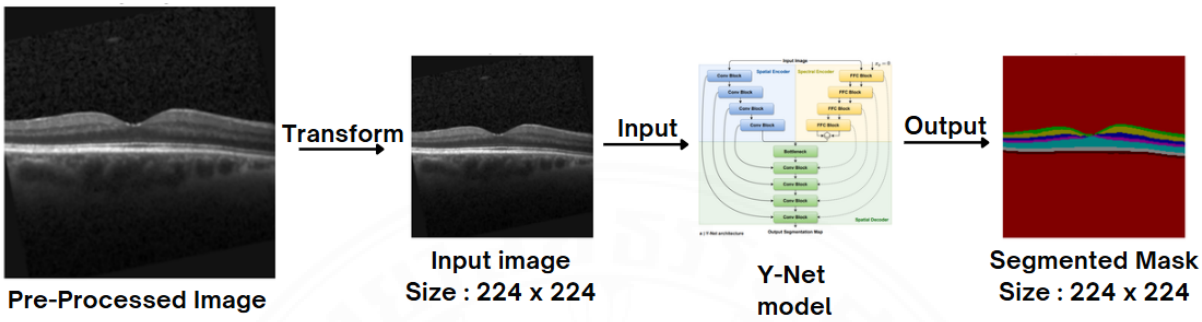

Figure 3.7 Illustration of Y-Net segmentation pipeline.

Prior to inference, each OCT image is transformed to conform to the model's input requirements. Specifically, images are resized to 224×224 pixels and normalized using a mean value of 46.3758 and a standard deviation of 53.9434. These normalization parameters are sourced from the original Y-Net implementation and ensure consistency with the model's training conditions.

Following pre-processing, images are fed into the Y-Net model, which produces pixel-wise segmentation maps delineating each of the seven retinal layers and the background. An example of the resulting segmentation output is shown in Figure 3.8.

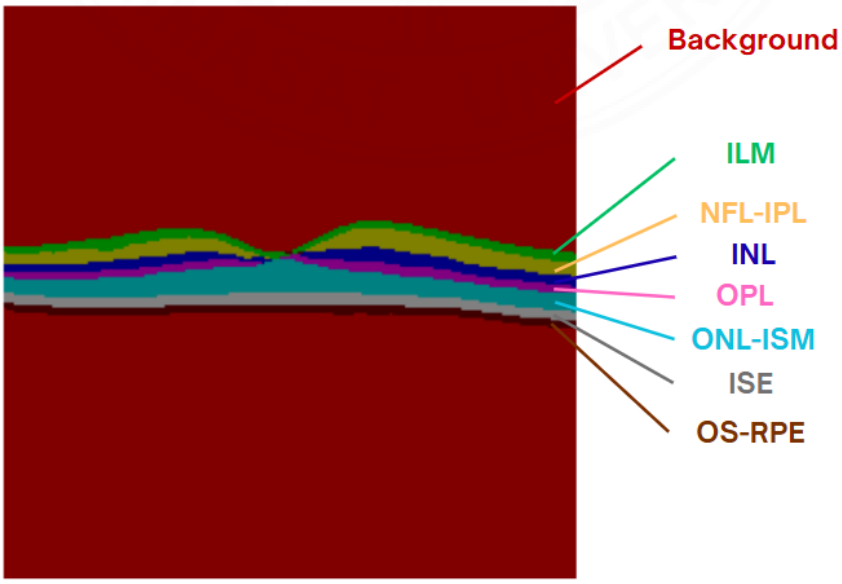

Figure 3.8 Illustration of segmented mask from Y-Net model.

3.2.3 Feature Extraction

After obtaining the segmented mask, we denote L as any individual segmented retinal layer, and G as the region encompassing all layers collectively. From these delineated regions, we extract both global and local features based on thickness and area measurements, as detailed below. This structured feature extraction enables comprehensive characterization of retinal morphology and facilitates analysis of inter-layer spatial relationships critical for classification and clinical interpretation. Figure 3.9 illustrates examples of the global piece G and local piece L.

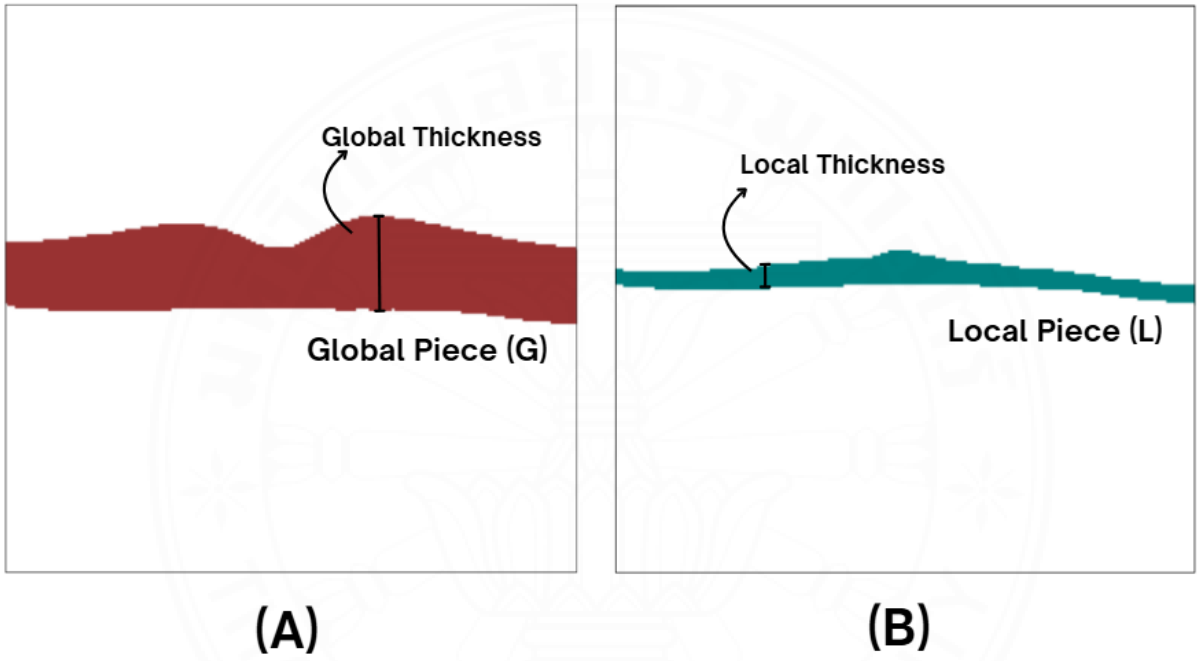

Figure 3.9 (A) Global piece G. (B) Local piece L.

Global Features (from combined retinal region G)

Nine features characterize overall retinal morphology:

1. **Thickest Layer (GTK):** Layer with the highest average thickness.
2. **Thinnest Layer (GTN):** Layer with the lowest average thickness.
3. **Largest Area Layer (GLA):** Layer occupying the largest pixel area.
4. **Smallest Area Layer (GSA):** Layer with the smallest pixel area.
5. **Area (GA):** Total pixel count of G.
6. **Maximum Thickness (GMaxT):** Maximum thickness value across all locations.

7. **Minimum Thickness (GMinT):** Minimum thickness value across all locations.

8. **Coefficient of Variation (GCV):** Variability measure of thickness values across G.

9. **Ratio of Maximum to Minimum Thickness (GMaxT-GMinT_Ratio):** Ratio between maximum and minimum thickness across the retina.

Local Features (from each individual layer L)

Five features are derived for each layer, making a total of thirty-five local features, including:

1. **Thickness Ratio at Global Maximum Thickness Location (LT-GMaxT_Ratio):** Ratio of the layer's thickness to the global maximum thickness at location x_{Gmax}.

2. **Thickness Ratio at Global Minimum Thickness Location (LT-GMinT_Ratio):** Ratio of the layer's thickness to the global minimum thickness at location x_{Gmin}.

3. **Maximum-to-Minimum Thickness Ratio within Layer (LMaxT-LMinT_Ratio):** Ratio of the maximum thickness to the minimum thickness within the layer L.

4. **Coefficient of Variation of Thickness (LCV):** Measure of the variability of thickness values across all locations within layer L.

5. **Relative Area Ratio (LA-GA_Ratio):** Ratio of the area of layer L to the total retinal area G.

Adjacent Layer Pair Features

For six pairs of adjacent layers, 12 features quantify inter-layer relationships:

1. **Average thickness ratio between adjacent layers (LAjT_Ratio):** This feature measures the ratio of the average thickness of the layer to that of its adjacent layer.

2. **Average area ratio between adjacent layers (LAjA_Ratio):** This feature measures the ratio of the area of the layer to that of its adjacent layer.

Combining global features, local features, and adjacent pair features yields a total of **56 features** extracted from the segmented masks, enabling detailed characterization of retinal morphology and inter-layer spatial relationships for improved classification and clinical insight.

3.3 Classification

To classify OCT images as normal or abnormal, we employed seven ML models: Decision Tree, Adaptive Boosting (AdaBoost), Light Gradient-Boosting Machine (LGBM), Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), and K-Nearest Neighbors (kNN). Each model was trained on the LG and LSR separately and evaluated based on same metrics.

CHAPTER 4
EXPERIMENTS

4.1 OCT Dataset Description

This study utilized a privately curated subset of 1,200 OCT images, equally divided between normal (600) and abnormal (600) cases. The abnormal category includes three retinal diseases—CNV, DME, and drusen—each represented equally. The images were selected from the publicly available dataset provided by Kermany et al. (2018).

The dataset was split into a training set of 1,000 images (500 normal and 500 abnormal) and a testing set of 200 images (100 normal and 100 abnormal). Additionally, 5% of the training set (i.e., 50 images) was randomly selected to serve as the validation set. This dataset configuration is consistent with the prior work by Numsonthi et al. (2025), enabling a direct performance comparison.

4.2 Implementation Details

All feature extraction and classification processes were implemented in Python. Hyperparameters for each model were optimized using grid search based on performance on the validation set.

4.3 Evaluation Metrics

The evaluation employed five standard classification metrics: Accuracy, Precision, Recall, F1-Score, and False Negative Rate (FNR). These metrics are mathematically defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

$$\text{FNR} = \frac{FN}{TP + FN} = 1 - \text{Recall} \tag{4.5}$$

Here, TP, TN, FP, and FN refer to true positives, true negatives, false positives, and false negatives, respectively.

CHAPTER 5
RESULTS AND DISCUSSION

5.1 Classification Performance

5.1.1 Performance of LG

The confusion matrices for each model applied to LG are shown in Figure 5.1. These matrices summarize classification outcomes in terms of TP, TN, FP, and FN.

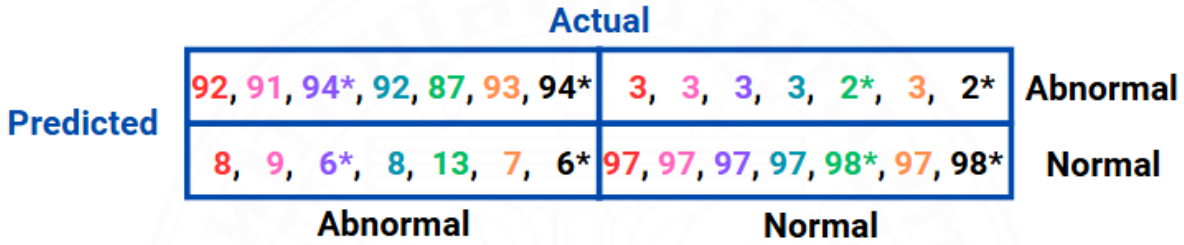

Figure 5.1 Confusion metrics report of LG for classification. Confusion matrices of Logistic Regression (Red), AdaBoost (Pink), Decision Tree (Purple), Random Forest (Blue), kNN (Green), SVM (Orange), and LGBM (Black). Asterisks denote best-performing metrics in each category.

Among all classifiers, the LGBM model demonstrated the strongest performance, achieving the best values for TP, TN, FP, and FN. The Decision Tree model produced similar TN and FP values, while kNN matched LGBM in TP and FN. Depending on the evaluation objective (e.g., minimizing false positives or maximizing true positives), Decision Tree and kNN remain viable alternatives.

5.1.2 Performance of LSR

The confusion matrices for each model applied to LSR features are shown in Figure 5.2. These matrices summarize classification outcomes in terms of TP, TN, FP, and FN.

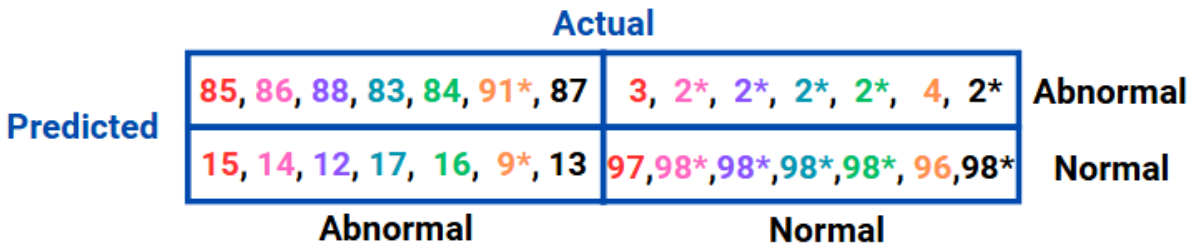

Figure 5.2 Confusion metrics report of LSR for classification. Confusion matrices of Logistic Regression (Red), AdaBoost (Pink), Decision Tree (Purple), Random Forest (Blue), kNN (Green), SVM (Orange), and LGBM (Black). Asterisks denote best-performing metrics in each category.

Among all classifiers, the LGBM model achieved the best values for TN and FP, while SVM achieved the best values for TP and FN. However, in terms of accuracy and FNR, the SVM model achieved the highest accuracy of 93.5% with the lowest FNR among all classifiers of 0.09%.

5.1.3 Performance Comparison

Figure 5.3 compares the proposed models: LG and LSR against the state-of-the-art method that uses line profile features (LP) proposed by Numsonthi et al. (2025).

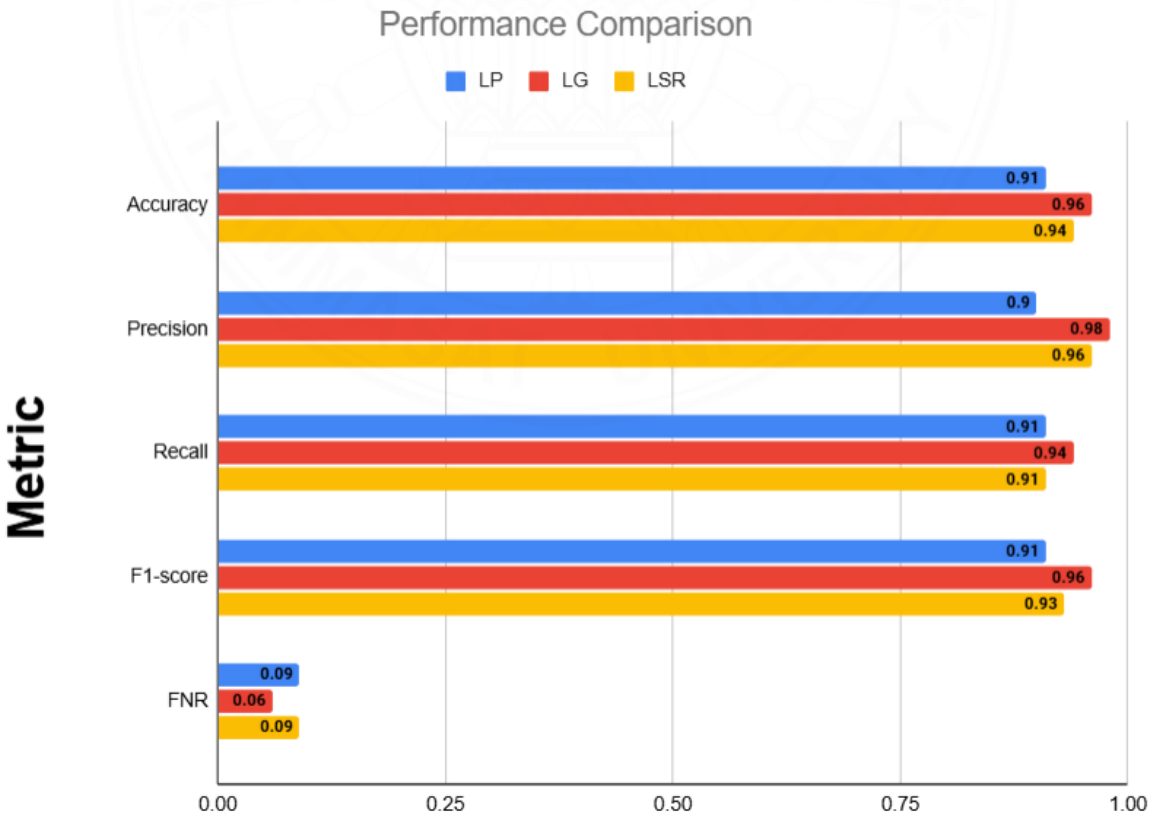

Figure 5.3 Metric comparison between MLLG and MLLP.

As shown, LG achieved superior results across all categories. LG achieved the highest accuracy at 96%, followed by LSR at 94% and MLLP at 91%. Precision is also highest for LG at 98%, surpassing LSR (96%) and MLLP (90%). Recall values are equal for MLLP and LSR at 91%, while LG is slightly higher at 94%. For the F1-score, LG leads with 96%, followed by LSR at 93% and MLLP at 91%. The false negative rate (FNR) is lowest for LG at 6%, indicating fewer missed positive cases, while both MLLP and LSR share a higher FNR of 9%. Overall, LG demonstrates superior performance, particularly in precision, accuracy, F1-score, and FNR, highlighting its effectiveness relative to MLLP and LSR.

5.2 Discussion

The results demonstrate the effectiveness of using structural and geometric features from retinal layers for automated OCT image classification. Two feature sets were examined: high-level LG from the ILM and RPE layers, and detailed LSR from seven segmented retinal layers.

The first study, LG with four features, yielded exceptional classification performance. The best-performing model in this study, LGBM, achieved an accuracy of 96%, precision of 98%, and a low false negative rate of 6%.

The second study, LSR, expanded the analysis to a granular level by extracting 56 features from seven segmented retinal layers, capturing thickness, area, and inter-layer dynamics. While this approach was designed to provide deeper clinical insights, its best-performing model in this study, SVM, achieved a slightly lower accuracy of 93.5% and a higher FNR of 9% compared to the LG method. This counterintuitive result, where more comprehensive features led to slightly lower binary classification performance, may be attributable to several factors:

Potential factors for performance discrepancy between LG and LSR

1. **Lack of pathology-focused features:** While the LG features (foveal concavity, bilateral symmetry, RPE smoothness) were chosen because they directly represent the fundamental structural integrity of a healthy retina. A healthy fovea has a distinct concave shape, and the retinal layers are symmetrical and smooth. Pathological conditions like macular edema, drusen, or CNV often cause direct and obvious disruptions to these specific high-level characteristics (e.g., flattening the fovea, breaking symmetry, creating bumps in the RPE layer). Therefore, these few features are highly ”focused” and act as powerful, direct indicators for the binary classification task of healthy vs. unhealthy.

   On the other hand, LSR features are broader. The approach extracts 56 features, capturing a vast amount of granular data, including thickness, area, and various ratios for seven

different layers without specific focus on particular layers.

2. **Feature redundancy and dimensionality:** The large set of LSR features may contain redundant or less discriminative information, potentially complicating the decision boundary for the classifiers. While these features can potentially offer deeper clinical insights into how a retina is deteriorating, they also risk introducing noise and irrelevant information for a simple binary classification.

3. **Segmentation dependency:** Performance of LSR is heavily reliant on the accuracy of the Y-Net segmentation model. In cases of severe pathology where retinal layers are highly distorted, segmentation errors could introduce noise into the feature set, degrading classifier performance.

The high performance of the LG approach, despite its relative simplicity, underscores its potential for practical clinical applications, such as rapid pre-screening. The features are not only computationally simple to extract but are also clinically intuitive, addressing the black-box problem often associated with deep learning models.

Given the limitations and challenges of LSR, future work could focus on addressing the identified limitations and synergizing the strengths of both methodologies. A key direction would be to apply feature selection techniques to the comprehensive LSR set to isolate the most impactful predictors, thereby reducing noise and redundancy. Additionally, exploring alternative, more robust segmentation algorithms could enhance the accuracy of layer delineation, particularly in severely diseased images where current methods may falter. A powerful next step would be to create a hybrid model that integrates the focused, highly discriminative LG with a refined subset of the most informative LSR features.

It is also important to note that the dataset in this study includes common retinal diseases such as AMD, DME, and drusen, which have high prevalence. However, it lacks representation of a broader range of retinal diseases found in real life. Although the proposed feature sets perform well on this dataset, this may reflect its adaptation to the specific diseases included. To ensure broader generalizability, the method should be further validated using a more diverse dataset.

CHAPTER 6
CONCLUSION

This thesis introduced and evaluated two complementary machine learning frameworks for the automated classification of retinal OCT images, demonstrating that analyzing retinal layer characteristics provides a highly effective basis for automated disease detection.

The first study, called Layer Geometry (LG), developed a framework using simple, clinically intuitive geometric features from the ILM and RPE layers, such as foveal concavity, bilateral symmetry, and RPE smoothness. This approach achieved an outstanding accuracy of 96% and a low false negative rate of 6% using a Light Gradient Boosting Machine (LGBM) model. It significantly outperformed a baseline method that used intensity-based line profiles, highlighting the robustness of structural biomarkers. This method is computationally efficient and highly accurate, making it suitable for rapid pre-screening applications.

The second study, called Layer Structural and Relational (LSR), expanded this analysis by extracting a comprehensive set of 56 features derived from the thickness, area, and inter-layer relationships of seven segmented retinal layers. This deeper, layer-wise analysis yielded slightly lower accuracy for binary classification, achieving 94% accuracy and a slightly higher FNR of 9%. This approach revealed that retinal layers exhibit distinct patterns and that retinal deterioration can be captured through changes within the retinal structure.

In conclusion, this research successfully demonstrates that machine learning models based on retinal layer morphology from simple geometries to complex inter-layer dynamics, offer a powerful, transparent, and accurate framework. Both frameworks outperformed the state-of-the-art approach by 5% and 3%, respectively. The proposed frameworks provide an alternative to black-box deep learning methods and support ophthalmologists by improving diagnostic efficiency and providing deeper clinical insights, contributing to the early detection and prevention of vision loss.

Future work will aim to improve classification performance by integrating all features, applying feature importance techniques, and incorporating feature interpretability to enhance understanding. Additionally, experiments will be conducted on more diverse datasets to further validate the approach and its generalizability.

REFERENCES

Abhishek, A., Berendschot, T., Rao, S., & Dabir, S. (2014). Segmentation and analysis of retinal layers (ilm & rpe) in optical coherence tomography images with edema. *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, 204–209.

Chanpipattanachai, C., Potchongruk, O., & Empraneat, Y. (2026). Unhealthy oct image detection using machine learning based on line profile features [In press]. *The 11th International Conference on Digital Arts, Media and Technology (DAMT) and 9th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (NCON)*.

Farshad, A., Yeganeh, Y., Gehlbach, P. L., & Navab, N. (2022). Y-net: A spatiospectral dual-encoder network for medical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. https://api.semanticscholar.org/CorpusID:250334726

Gonzalez, R., & Woods, R. (2008). *Digital image processing* (3rd). Prentice Hall.

Hussain, M., Bhuiyan, A., Luu, C., Smith, R., Guymer, R., Ishikawa, H., Schuman, J., & Ramamohanarao, K. (2018). Classification of healthy and diseased retina using sd-oct imaging and random forest algorithm. *PLOS ONE*, *13*(6), 1–17. https://doi.org/10.1371/journal.pone.0198281

Kermany, D., Zhang, K., & Goldbaum, M. (2018). Labeled optical coherence tomography (oct) and chest x-ray images for classification [Accessed: 2025-01-15]. https://www.kaggle.com/datasets/paultimothymooney/kermany2018

Khalid, S., Akram, M., Hassan, T., Nasim, A., & Jameel, A. (2017). Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age related macular degeneration from optical coherence tomography images. *Biomed Research International*, *2017*, 7148245. https://doi.org/10.1155/2017/7148245

Numsonthi, A., Polpanich, B., Rawanggij, K., Jundang, K., Dorji, T., & Aimmanee, P. (2025). Rapid and accurate abnormal oct image detection using ml-based on intensity line profile features. *2025 17th International Conference on Knowledge and Smart Technology (KST)*. https://doi.org/10.1109/KST65016.2025.11003336

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (2007). *Numerical recipes 3rd edition: The art of scientific computing* (3rd). Cambridge University Press.

Salaheldin, A., Wahed, M., & Saleh, N. (2024). A hybrid model for the detection of retinal disorders using artificial intelligence techniques. *Biomedical Physics & Engineering Express*, *10*(5), 055005. https://doi.org/10.1088/2057-1976/ad4c5c

Syed, A., Hassan, T., Akram, M., Naz, S., & Khalid, S. (2016). Automated diagnosis of macular edema and central serous retinopathy through robust reconstruction of 3d retinal surfaces. *Computer Methods and Programs in Biomedicine*, *137*, 1–10. https://doi.org/10.1016/j.cmpb.2016.09.004

Szeliski, R. (2022). *Computer vision: Algorithms and applications*. Springer Nature.

Yousefi, J. (2011). Image binarization using otsu thresholding algorithm.

BIOGRAPHY

Name	Methawee Ratanapunperm
Education	2023: Bachelor of Engineering
	(Computer Engineering)
	Sirindhorn International Institute of Technology
	Thammasat University

Publication

Ratanapunperm, M., et al. (2025). Automated detection of unhealthy retinal optical coherence tomography images using geometrical features for machine learning modelling. *Proceedings of the International Conference on Information and Communication Technology for Intelligent Systems*, accepted March 2025.