Deep Learning-assisted Interpretable Retinopathy of Prematurity (ROP) Diagnosis

MS(Research) Thesis

By

Urvesh Trivedi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

July 2025

Deep Learning-assisted Interpretable Retinopathy of Prematurity (ROP) Diagnosis

A THESIS

submitted to the

INDIAN INSTITUTE OF TECHNOLOGY INDORE

in partial fulfillment of the requirements for

the award of the degree

of

MS(Research)

By

Urvesh Trivedi

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

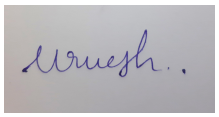
July 2025

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis, titled **Deep Learning-assisted Interpretable Retinopathy of Prematurity (ROP) Diagnosis**, is submitted in partial fulfillment of the requirements for the Master of Science (Research) degree in the Department of Computer Science and Engineering at the Indian Institute of Technology Indore. It presents the original work I conducted between August 2023 and April 2025.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.



24 / 02 / 2026

Signature of the Student with Date

(Urvesh Trivedi)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of Thesis Supervisor with Date

(Prof. Abhishek Srivastava)

Urvesh Trivedi has successfully given his MS(Research) Oral Examination held on 23 / 02 / 2026

Signature of Thesis Supervisor

Date: 25/ 02/ 2026

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to everyone who contributed significantly to this journey. First and foremost, I am deeply indebted to my supervisor, **Prof. Abhishek Srivastava**, for his unwavering inspiration throughout my research. This research work could not have been completed without his constant guidance and direction. His continuous support and encouragement have motivated me to remain streamlined in my research.

I am thankful to all my research committee members, for taking some valuable time to evaluate my progress throughout the course. Their good comments and suggestions helped me to improve my work at various stages. I am also grateful to **Dr. Ranveer Singh**, HOD of Computer Science and Engineering Department and **Dr. Somnath Dey**, former HOD of the Computer Science and Engineering Department, for their help and support.

My sincere acknowledgement and respect to **Prof. Suhas Joshi** , Director, Indian Institute of Technology Indore for providing me the opportunity to explore my research capabilities at Indian Institute of Technology Indore.

I want to express my heartfelt respect to my parents, siblings, and friends for the love, care, and support they have provided to me throughout my life.

Finally, I am thankful to all who directly or indirectly contributed, helped, and supported me.

Urvesh Trivedi

Abstract

Retinopathy of Prematurity (ROP) is a sight-threatening retinal eye disease primarily affecting premature babies due to the growth of abnormal blood vessels in the retina. Early detection and timely treatment is very important to prevent irreversible vision loss; however, effective screening for ROP is limited by the lack of resources and trained ophthalmologists, especially in underserved and rural areas. In recent years, several studies have been conducted on the development of reliable AI-based screening systems. However, due to the lack of well-annotated public datasets most of them gives partial diagnostic solutions and are limited to experimental research, using single-central datasets.Our work presents a deep learning-assisted diagnostic framework for automated and interpretable ROP screening. The proposed system is composed of three key modules: (1) zone separation, (2) ridge (or demarcation line) detection, and (3) blood vessel segmentation—each targeting clinically relevant retinal features defined by the International Classification of ROP (ICROP) guidelines. To support and evaluate the framework,we created a comprehensive and expert-annotated dataset - Macretina, consisting of 1,432 retinal fundus images from 112 premature infants. Finally, we integrated our novel diagnostic framework into a lightweight mobile application, designed for real-time deployment in neonatal care units. Our proposed solution demonstrates high diagnostic accuracy, interpretability, and scalability, offering a clinically viable tool for early ROP screening, especially in low-resource and telemedicine settings.

List of Publications

1. **Trivedi Urvesh**, Abhishek Srivastava, and Pratik Mahajan. *Macretina: a dataset, to support deep learning assisted retinopathy of prematurity diagnosis*, Scientific Reports, 2025 (Accepted).

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

ROP Retinopathy of Prematurity

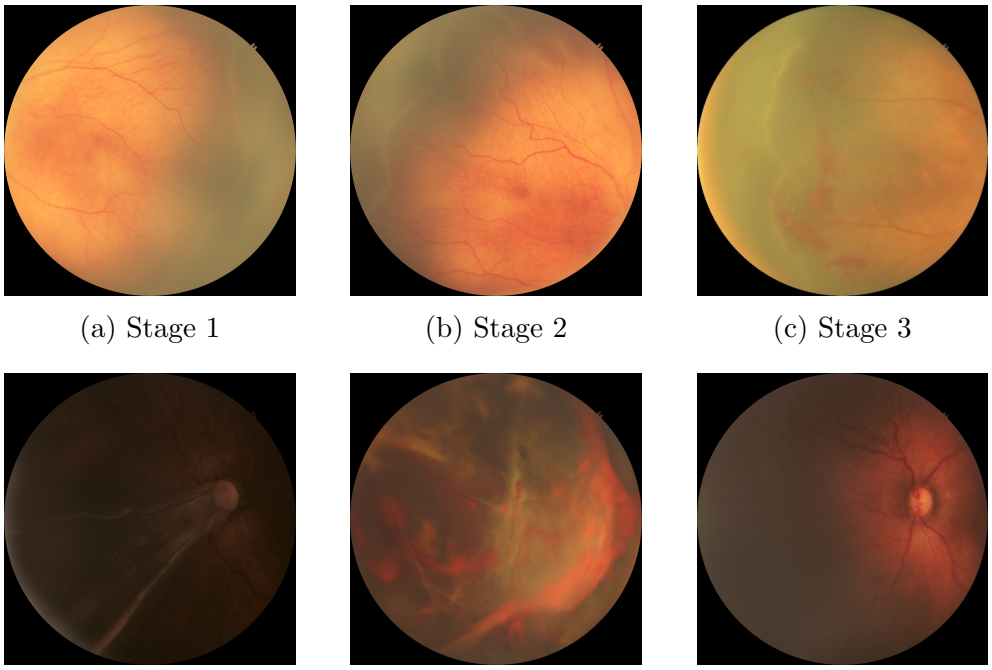
DCNN Deep Convolutional Neural Network

XAI Explainable AI

Chapter 1

Introduction

Premature babies are the babies which born before 31 weeks of pregnancy and have low birth weight (less than 3 pounds). Retinopathy of prematurity (ROP) is one of the most frequent and severe eye diseases found in premature babies. The main cause behind ROP is the growth of abnormal blood vessels in the retina (light-sensitive layer) of premature babies. The development of blood vessels in the eyes is usually completed a few weeks before birth. However, premature babies may face various factors, including medications, changes in temperature and oxygen-level in the atmosphere, and bright lighting, which may affect the development of the blood vessels in their eyes. ROP is sometimes mild and naturally cured without any treatment but in severe cases it requires early diagnosis followed by proper medical treatment to prevent visual impairment or blindness. According to [1, 2], India has the highest rate of preterm births, one of the highest in the world, contributing to the increased incidence of retinopathy of prematurity (ROP). In middle-income countries like India where more than 64% of population lives in rural areas or small villages [3], there is a lack of necessary resources and expertise for timely screening of preterm infants, leading to late diagnoses and poor treatment. Thus, accessibility and quality of nenotal care are the most affecting factors for the rise in cases of severe ROP incidences in india [4].



(a) Stage 1 (b) Stage 2 (c) Stage 3

(d) Stage 4 (e) Stage 5 (f) Aggressive ROP

Figure 1.1: Sample images from the Macretina dataset representing different stages of Retinopathy of Prematurity (ROP) as defined by ICROP standards.

The International Classification of Retinopathy of Prematurity (ICROP) guideline provides standard rules for the detection and classification of retinopathy of prematurity (ROP) into different stages and zones to measure its severity [5]. Zone-level classification help to locate the degree of vascularization, dividing the retina into three concentric areas centered around the optic disc, namely Zone-1, Zone-2, and Zone-3, with Zone-1 indicating the most severe and posteriorly located disease. Stage-level classification is described by the presence of retinal findings. In stage-1 (figure 1.1(a)), thin white line is found between vascular and avascular regions of retina, called demarcation line, which requires no treatement unless it progresses. In stage-2 (figure 1.1(b)), this demarcation line becomes thick and called ridge; laser therapy is required if progression occurs. In stage-3 (figure 1.1(c)), rapid growth of extraretinal neovascularization (formation of abnormal blood vessel) is found near the ridge, causing a ragged appearance near the ridge; often treated with laser photocoagulation or posterior laser treatment. In Stage-4 (figure 1.1(d)), partial retinal detachment begin at the ridge and may extends towards the fovea; treated with a combination of laser

therapy, injections, and surgery. In stage-5 (figure 1.1(e)), total retinal detachment is observed; has to be managed surgically but the successful anatomical results after surgery are only seen in 20%–50% of cases [6]. In a recent update, a more severe form of the disease was introduce called Aggressive ROP (A-ROP) (figure 1.1(f)), a rapidly progressing form of ROP commonly seen in Zone-1 or posterior Zone-2, characterized by severe vascular abnormalities without the typical ridge formation. Babies with A-ROP require urgent treatment as it can quickly lead to total retinal detachment. Thus, it is necessary to stop the progression of ROP at early stage with early diagnosis, followed by a proper medical treatment. Lack of ROP experts and nenotal care units, especially in a country like India (with more rural areas), calls the need for an automated diagnosis system for ROP.

In recent years, several studies have been conducted on AI-based retinopathy of prematurity (ROP) diagnosis, using deep learning (DL) techniques for automated detection and classification of pathological features in retinal fundus images. Such studies can be helpful in building AI-based screening systems that can assist ophthalmologists in identifying key retinal features, such as growth of abnormal blood vessel, affected zone, and presence of demarcation line/ridge, which are essential for ROP diagnosis and severity analysis. However, existing AI-driven studies often lack interpretability, miss complex and severe cases of ROP such as A-ROP, include inconsistent zone separation logic and do not provide a complete framework for ROP diagnosis. Some studies have also highlighted the lack of well-annotated public datasets for AI-based ROP diagnosis [7, 8, 9, 10], which is crucial to develop reliable deep learning models for accurate diagnosis. Moreover, existing public datasets have significant limitations—some datasets only provide ground truth annotations for a single pathological feature, making them ineffective for a complete ROP diagnosis and severity analysis [11, 12, 13], while datasets covering annotations for multiple pathological features, alone are not sufficient to effectively train deep learning models [14].

To address these challenges, we propose a deep learning-assisted comprehensive framework for automated ROP diagnosis. The framework is mainly composed of three submodules: 1) zone separation, 2) ridge (or demarcation line) detection, and 3)

blood vessel segmentation—each targeting a pathologically relevant feature for ROP screening. These modules are validated on real-world data and integrated into a mobile application, making the solution deployable in real-time clinical environments, particularly in low-resource settings. To effectively train and validate each component of our proposed framework, we also introduce **Macretina**, a comprehensive and expert-annotated ROP dataset consisting of retinal fundus images of premature babies captured using the Neo imaging system. To prepare this dataset, we collected 1,432 retinal fundus images of 112 premature babies from a local hospital. The dataset was developed under the guidance of ROP experts and is specifically designed to evaluate deep learning models used across different submodules of our proposed framework.

The key contributions of our work are summarized as follows:

- We present a novel ROP dataset consisting of retinal fundus images of premature babies, collected from a local hospital, and annotated with the help of ROP experts to ensure reliability and clinical relevance.
- We propose a deep learning-assisted diagnostic framework designed to perform a complete and automated ROP diagnosis.
- We also introduce a novel ROP zone mapping logic based on the detected Optic Disc location, enabling consistent and interpretable zone separation following the ICROP guidelines.
- For ridge (or demarcation line) detection, we design a binary classification module, validated using Explainable AI (XAI) technique to enhance interpretability and transparency in clinical decision support.
- For Blood Vessel segmentation, we evaluate widely used U-Net based models on our dataset to test their generalizability and performance on Aggressive ROP (A-ROP) cases covered in the Macretina dataset.
- All three modules are integrated into a lightweight mobile application designed for real-time deployment in neonatal care units, aiming to support ophthalmologists during initial screening.

Chapter 2

Literature Review

Retinopathy of Prematurity (ROP) is one of the primary causes of childhood blindness. Therefore, early detection is very crucial for timely intervention in order to prevent vision loss, as the disease progresses rapidly and unpredictably in some cases. Fundus imaging is widely used for ROP screening, as it provides a clear view of the retinal structures, helping doctors to identify pathological changes at an early stage. Given the importance of early and accurate diagnosis, significant progress has been made over the years to improve the screening process. Thus, we have divided this section into three subsections, highlighting key approaches used by prior researches, followed by the available opportunities or gaps in existing research.

The subsections are mentioned below :

- Manual Screening
- Computer-Aided Diagnosis (CAD)
- Deep Learning Based Approaches
- Opportunities or Gaps in Existing Research

2.1 Manual Screening

Manual screening is one of the oldest and most widely used approaches for Retinopathy of Prematurity (ROP) diagnosis, where a ROP specialist examines reti-

nal fundus images to identify signs of the disease. Although this method is clinically significant, it is time-consuming and prone to subjectivity at the same time [15]. The results of such manual screening are based on subtle cues such as color, texture, vessel tortuosity, and retinal structure, making it highly dependent on the expertise and experience of the clinician. As a result, this often leads to inconsistencies and variability in diagnostic outcomes. Moreover, there is a global shortage of ophthalmologists trained for ROP screening, which further reduces the effectiveness and accessibility of ROP screening programs, especially in developing countries like India, with increasing premature births [16, 17].

2.2 Computer-Aided Diagnosis (CAD)

Computer-Aided Diagnosis (CAD) system is a viable alternative solution for analyzing medical images in retinopathy of prematurity (ROP) screening. CAD tools are designed to help clinicians in screening, monitoring, and diagnosing retinal diseases by allowing quicker, more accurate, and consistent decision-making processes [18]. Unlike fully automated diagnostic algorithms that provide a final diagnosis independently, CAD systems serve as supportive tool instead of providing a clinical decision. Despite their potential help in the screening process, traditional CAD approaches often struggle in understanding the progression of complex retinal diseases such as diabetic retinopathy, glaucoma, age-related macular degeneration (AMD), ROP, and Plus disease [19, 20, 21]. This is especially true in case of infant fundus images, as they are typically difficult to process due to low image quality, poor contrast, and overlapping anatomical structures. To address these challenges, machine learning (ML), and more recently deep learning (DL) techniques, have been increasingly adopted to automate the ROP screening process[22].

2.3 Deep Learning Based Approaches

The introduction of artificial intelligence (AI), particularly deep learning algorithms, has significantly improved the detection and diagnosis of Retinopathy of Prematurity (ROP). These algorithms have demonstrated remarkable diagnostic performance, with sensitivity rates averaging 95.72% and specificity reaching 98.15%, thereby showing capabilities comparable to ROP experts [23]. In addition, AI-based ROP screening systems play a very important role in addressing challenges due to limited access to qualified professionals, especially in countries such as India, where the burden of ROP is increasing due to the high rate of preterm births [24].

Kumar et al.[8] proposed a Deep Convolutional Neural Network (DCNN)-based approach for ROP diagnosis, integrating YOLOv5 for optic disc detection and Pix2Pix & U-Net for blood vessel segmentation. Their proposed architecture is trained on public retinal fundus datasets and validated on a small ROP dataset collected from a local hospital, achieved 98.94% accuracy for OD detection and a Dice coefficient of 0.455–0.528 for BV segmentation on ROP dataset. Their system accurately diagnosed ROP in Zone-1 with 88.23% accuracy, highlighting its potential for automated screening, especially in low-resource settings. However, the study points out the limitation of training with small datasets, which can affect the generalizability of the model in diverse populations, suggesting the need for larger datasets.

In Agrawal et al.[7], segmentation of retinal structures is defined for the accurate identification of features related to ROP diagnosis. The challenges involved in the segmentation task, such as the need for precise annotations and the unavailability of complete datasets, underscore the importance of this research. To solve these difficulties, the authors developed two primary datasets (HVDROPDB-BV and HVDROPDB-RIDGE) [14] and employed U-Net architecture along with its variation that includes squeeze & excitation (SE) blocks and attention gates (AG). These enhancements achieved promising performance metrics, with the AG U-Net achieving a sensitivity of 96% and a specificity of 89% for detecting ROP stages through test images.

Another study by Agrawal et al.[25] introduces a novel deep learning framework for classifying retinal fundus images into normal and ROP stages - 1, 2, and 3. It uses an attention gate U-Net for demarcation line segmentation and the Random Forest (RF) algorithm for effective feature selection. Their proposed model gives better performance compared to VGG19 and achieves competitive results with Xception and InceptionV3 on the HVDROPDB[14] dataset. Although the study achieved very good results in identifying early-stage ROP, the classification accuracy for stages 2 and 3 can be improved further. Furthermore, the framework gave a significantly smaller model size and lower computational complexity, making it suitable for the integration into real-time fundus imaging systems. The study also highlights the importance of explainable AI in medical imaging by providing localization of the demarcation line, helping ophthalmologists in early ROP detection and intervention.

The study by Wang et al.[26] focuses on improving the segmentation of infant retinal images and providing a quantitative vascular analysis to enhance the diagnosis of Plus disease in Retinopathy of Prematurity (ROP). The authors developed two modified U-Net architectures: U-Net3 with dual attention modules for blood vessel segmentation and U-Net1 with reduced channels for optic disc segmentation. Their method was trained on a dataset of 40 manually annotated vessel segmentation images and 169 optic disc segmentation images, achieving an F1 score of 0.8116 and sensitivity of 0.8273 for vessel segmentation, while the optic disc segmentation attained an F1 score of 0.9346 and sensitivity of 0.9395. Additionally, they introduced a quantitative analysis of vascular features such as tortuosity, density, and width by defining an optimal region of interest (ROI) centered around the optic disc with a radius four times its diameter.

Omneya Attallah [27] introduces an automated deep learning-based diagnostic tool for Retinopathy of Prematurity (ROP) called DIAROP. Which uses transfer learning from four pre-trained Convolutional Neural Networks (CNNs) to extract spatial features, followed by the application of Fast Walsh Hadamard Transform (FWHT) for feature integration. The study evaluates three settings namely spatial feature extraction from CNNs, spatial-spectral feature extraction using FWHT, and integration of

the best spatial-spectral features. The results indicate that including spatial-spectral features improves diagnostic accuracy, with DIAROP achieving an accuracy of 93.2% and an AUC of 0.98. While DIAROP significantly reduces manual effort and examination time, the study acknowledges limitations such as the lack of segmentation techniques and severity classification. Future work aims to address these gaps by incorporating additional CNNs and expanding DIAROP's capabilities to detect and categorize different ROP severity levels.

Almeida et al.[28] presents a novel automated pipeline for ROP diagnosis which uses advanced image processing techniques such as CIELAB Enhancement, vesselness filtering, and a novel Optical Disc artifact removal approach. Their pipeline was evaluated on fundus images from Clarity RetCam3, Phoenix ICON, and the DRIVE database, achieving accuracy scores of 0.94, 0.94, and 0.95, respectively. Additionally, classification using DenseNet121 achieved an accuracy of 0.946, demonstrating the potential for fully automated ROP diagnosis.

The study by Liu et al.[29] gives an interpretable (XAI) machine learning approach to predict the risk of ROP in preterm infants using clinical data from 642 preterm infants. In this study six machine learning models were evaluated, with XGBoost achieving the best performance, obtaining an AUC 0.949 on the validation set. Key predictive features highlighted using SHapley Additive exPlanations (SHAP) include severe pre-eclampsia, Apgar score at 1 minute, gestational age at birth, very low birth weight, blood transfusion, and neonatal hyperglycemia. These findings highlight the model's potential for clinical application in early ROP risk prediction.

Wang et al.[30] presents a deep learning–based platform for automated ROP screening using retinal funuds images of preterm infant. When tested on 52,249 images, the platform achieved an AUC of 0.983 to 0.998 for individual classifiers, with a Cohen k of 0.86 to 0.98, comparable to ROP experts (0.93 to 0.98). CAM and DeepSHAP (XAI) was used in this study for model interpretability, generating heatmaps to explain predictions. This findings suggest the system's potential for routine ROP screening in hospitals, supporting early diagnosis and intervention. However, the study does not include zone level classification given by ICROP standards. Also, the Heatmaps given

by DeepSHAP mehtod for model interpretation are fragile and do not fully satisfy sensitivity and implementation invariance.

2.4 Opportunities or Gaps in Existing Research

Prior researches indicates the usage of deep learning models for automated ROP diagnosis. The gaps that were identified are mentioned below :

- Lack of ROP Datasets: The lack of publicly available, well-annotated, and diverse ROP datasets makes it difficult to train and develop robust AI models for ROP diagnosis.
- Lack of Interpretability: Many existing AI based studies for ROP diagnosis lack interpretability, making it difficult for ophthalmologists to trust and validate the automated findings.
- Inadequate handling of severe ROP cases: Current approaches often miss aggressive forms of the disease called Aggressive ROP (A-ROP), limiting their clinical utility.
- Inconsistent or absent zone mapping logic: Only few studies provides detailed and consistent methods for ROP zone identification, despite its critical role in severity analysis of the disease.
- Partial diagnostic frameworks: Most AI-based solutions focus only on isolated tasks (e.g., classification or segmentation) instead of providing an complete integrated framework covering all key ROP diagnostic features.

Chapter 3

Deep Learning-assisted Interpretable Retinopathy of Prematurity (ROP) Diagnosis

This chapter discusses in detail the proposed methodology for Deep Learning-assisted Interpretable Retinopathy of Prematurity (ROP) Diagnosis. The problem statement is clearly stated in Section 3.1; this is followed by the proposed methodology in Section 3.2. Subsequent sections provide step wise descriptions of the proposed methodology in detail.

3.1 Problem Statement

The problem of ROP diagnosis is framed as a multi-module deep learning task. Given a dataset D consisting of N retinal fundus images, represented by $I = \{I_i\}_{i=1}^{N}$. For each image I_i, there are three associated tasks: zone mapping $Z_i \in \{\text{I, II, III}\}$, ridge detection $R_i \in \{0, 1\}$ indicating the absence or presence of a demarcation line, and blood vessel segmentation $S_i = \{s_{ij}\}_{j=1}^{M}$, where $s_{ij} \in \{0, 1\}$ denotes pixel-wise vessel labeling. The objective is to build an interpretable and accurate deep learning system capable of jointly performing Z_i, R_i, and S_i for each I_i, suitable for real-time mobile deployment in ROP screening.

3.2 Proposed Methodology

The detailed architecture of the proposed approach is described in figure 3.1. The four main components are retinal imaging, image pre-processing, feature extraction, and ROP classification. The retinal fundus images of premature babies contains detailed information about the retinal structures. However, the collected retinal fundus images are unclear and contain noise, thus a pre-processing step is required to improve the quality of these images. The pre-processed images are then used to extract specific retinal features related to ROP which includes ridge/demarcation line, affected ROP zone, and blood vessel map. Finally, the classification unit combines the information from different feature extraction sub-units and gives visual representation of ROP screening through mobile application.

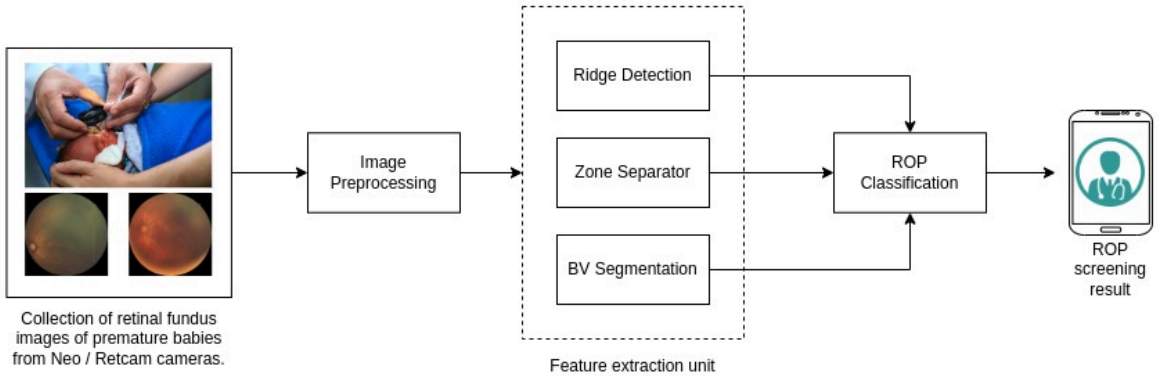


Figure 3.1: Detailed architecture of proposed methodology

3.2.1 Dataset Preparation

We collected 1432 retinal fundus images from 112 premature babies from the Macretina Hospital,Indore, India. These images (of size 2040 × 2040 pixels) were captured using the Neo-imaging system [31]. On average, 4 to 12 images were taken (from both eyes) per baby during regular retinal screening of premature babies at Macretina Hospital for ROP diagnosis. Before imaging, pupil dilation was done using diluted phenylephrine and tropicamide, given three times at 10-minute intervals. Feeding was also stopped after the final dose to meet optimal imaging conditions. All

personal details about the babies were kept private by the hospital during the dataset collection process. The collected fundus images covered different retinal views, such as posterior, temporal, superior, inferior, and nasal. However, we selected mainly temporal and posterior view retinal images, as they provide the most critical information about disease progression.

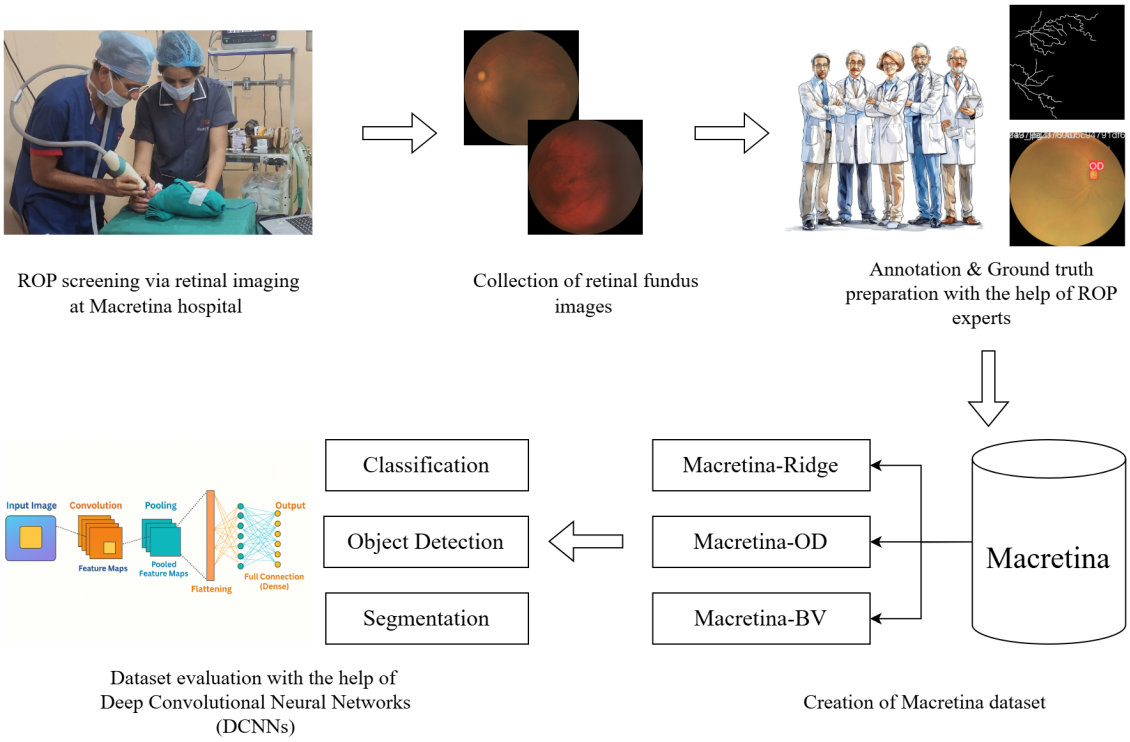


Figure 3.2: Dataset Preparation

One of the authors, Pratik Mahajan, a senior ophthalmologist at Macretina Eye Hospital with over 15 years of experience in retinal surgeries and pediatric ophthalmology, guided the annotation process with his clinical expertise. To gain a deep understanding of ROP pathology, the authors reviewed existing AI-based research in ROP screening in order to identify the key gaps and challenges present in recent studies. They also had detailed discussions with ROP experts to understand their clinical needs to prepare the dataset accordingly. As part of the process, the authors learned about Aggressive ROP (A-ROP), which is increasingly reported in India [4, 32]. They also gained practical exposure to different ROP stages (Stages 1–5) and learned why temporal and posterior views are essential in ROP diagnosis. Thus, the annotation

process was not only about marking the pathological features but also served as a valuable learning experience for the authors.

After removing noisy and irrelevant images, a total of 1,129 images were used for the labelling and annotation process. Of these, 250 images represent Stage 1, 2, and 3 ROP cases, around 200 images covers Aggressive ROP (A-ROP) cases, and 215 images are of normal healthy retina. This distribution ensures a diverse representation of ROP severity within the dataset. Recent studies are mainly focused on detecting retinal structures critical for ROP diagnosis—such as the optic disc, blood vessels, and the demarcation line/ridge—which are essential for identifying ROP zones and stages. Therefore, images with laser scars (due to laser therapy) and those showing partial or complete retinal detachment (Stage 4 and Stage 5 cases) were excluded. These cases were limited in number and not the main focus of our work. Moreover, these advanced stages involve significant retinal detachment, often reducing the visibility of key retinal structures, which makes them less suitable for AI-based ROP diagnosis. Hence, the Macretina dataset was specifically prepared to support the detection of pathological features in early-stage ROP cases.

The **Macretina** dataset is named after Macretina Hospital, Indore, India. It consists of posterior and temporal view retinal fundus images of premature babies with gestational age between 24-36 weeks and birth weight less than or equal to 3000 gm. All images are stored in JPEG format, with a resolution of 2040 × 2040 pixels and a file size between 150 to 250 KB. It includes three sub-datasets, presented in table 3.1, each designed for a specific task in AI-assisted ROP diagnosis. The **Macretina** dataset is publicaly available on Figshare.

Sub-dataset	**Application**	**Size**	**Label**
Macretina-Ridge	Classification	465	Two class (Ridge & No-Ridge)
Macretina-OD	Object detection	500	Bounding box
Macretina-BV	Segmentation	30	Blood vessel map

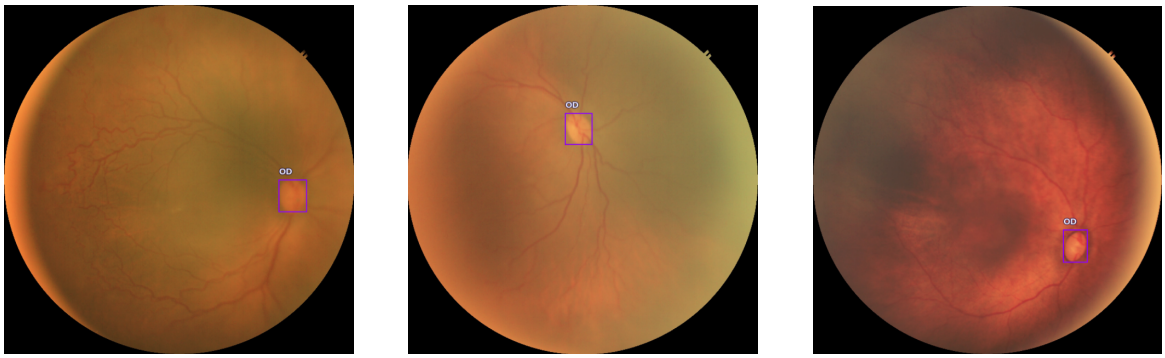
Table 3.1: Macretina Dataset

The **Macretina-Ridge** dataset is designed for binary classification, which can be helpful for training deep learning models to detect the presence of a ridge or demarca-

tion line in retinal fundus images. Identifying this feature at the right time is crucial to monitor ROP staging, as it can progress to a more severe disease state. It is one of the earliest and most significant indicators of disease progression, therefore its accurate detection is very important for early diagnosis and treatment. Sometimes it is difficult to find ridge or demarcation line in the retinal images of premature babies. Unlike other retinal structures, the appearance of a ridge or demarcation line varies depending on factors such as image quality, severity of the disease, and the retinal development stage. In some cases, it appears as a well-defined elevated structure, while in others, it is thin, blurry, and hardly visible. Additionally, imaging conditions—such as poor illumination, motion artifacts, and reflections from the vitreous (clear gel that fills the space between the lens and the retina) can hide the ridge, which makes manual labelling difficult.

To prepare accurate annotations, each image in the Macretina-Ridge dataset was manually labelled along with the team of ROP experts. Their expertise was required in distinguishing true ridge formations from imaging artifacts, to make sure that only clinically significant features were marked. This sub-dataset includes 465 images, with 250 images showing a ridge or demarcation line and 215 images representing normal retinal structures. By providing expert-labeled images for ridge/demarcation line detection, the Macretina-Ridge dataset serves as a valuable resource for developing AI models that can detect early signs of ROP progression. Such models can support ophthalmologists by differentiating normal and pathological retinal structures, improving early-stage ROP diagnosis, and helping in timely medical intervention for affected infants.

The **Macretina-OD** dataset is designed for optic disc (OD) detection, a crucial component in zone-based ROP analysis. It contains 500 retinal fundus images, each with a manually annotated rectangular bounding box around the optic disc as shown in figure 3.3. These annotations serve as ground truth for training Object Detection models to accurately locate the optic disc in unseen retinal images. Annotating the optic disc in infant retinal images is more challenging than in adults. In adults, the optic disc is larger, well-defined, and has clearly visible boundaries. In contrast,



(a) (b) (c)

Figure 3.3: Manually annotated bounding box around the optic disc (OD)

infant retinal images often show a smaller, blurry, and less distinct optic disc due to an underdeveloped retina and lower image contrast. Imaging conditions such as uneven illumination and a limited field of view further reduce its visibility. The small size and unclear boundaries of the optic disc in neonatal eyes make it difficult to draw precise bounding boxes for accurate localization. However, it is important to accurately compute the center and diameter of the optic disc for zone-based analysis.

To overcome these challenges, the Macretina-OD dataset includes accurate bounding box annotations, so that the extracted coordinates can be effectively used to compute the center and diameter of the optic disc which can be further used for zone-based ROP classification. In creation of this sub-dataset, mostly posterior (central) view retinal images were included, as they provide the clear and complete view of the optic disc. Since the coordinates of optic disc serves as a key reference for assessing disease progression, this sub-dataset provides valuable data for AI models that assist in ROP severity grading. By providing well-labeled posterior view images, the Macretina-OD dataset supports the development of AI models that can help clinicians in zonally classifying ROP cases to check disease severity and progression.

The **Macretina-BV** dataset is designed for blood vessel segmentation, which plays a very important role in ROP severity analysis. Since abnormal vascular development is a key indicator of disease progression, accurate segmentation is essential for AI-assisted ROP diagnosis. This sub-dataset includes 30 high-quality retinal fundus images, each paired with a manually created blood vessel mask as shown in figure 3.4.

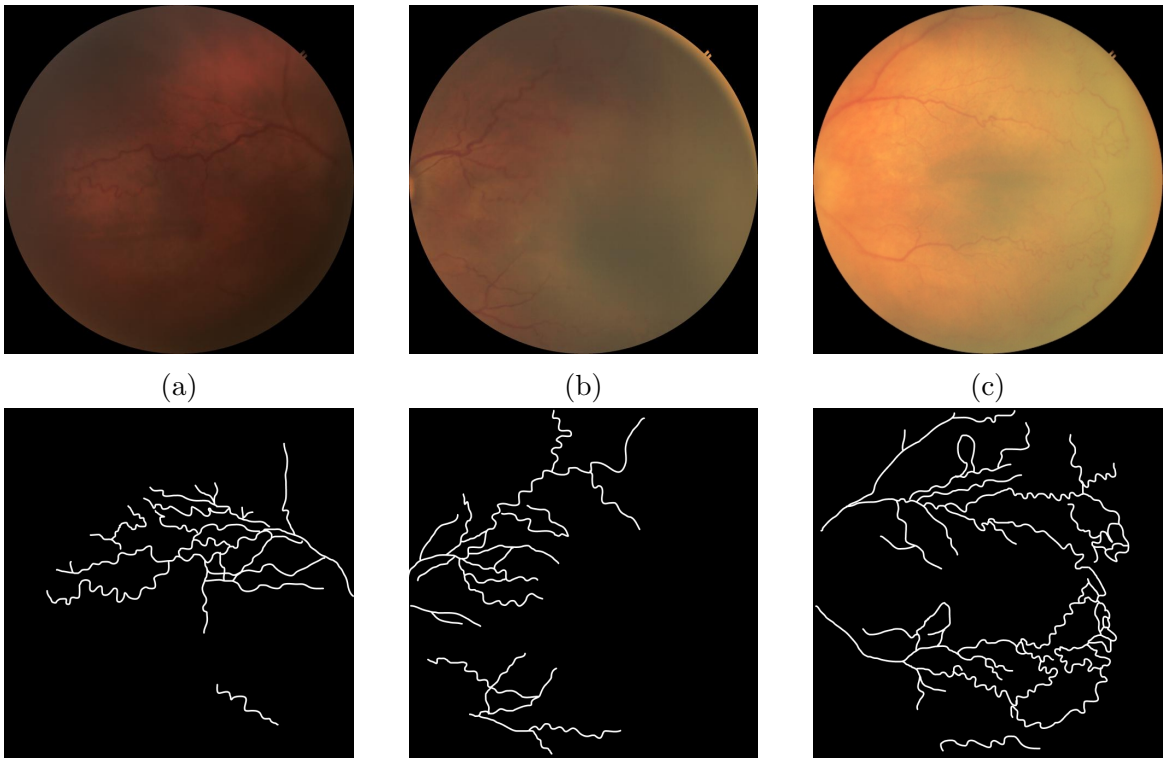


Figure 3.4: Neo Images (a,b,c) with their BV-Masks (e,f,g) from Macretina-BV dataset

These ground truths (blood vessel mask) were manually created using a Figma design tool with the help of a digital designer and a team of ROP experts. Due to the high cost and time-consuming nature of manual vessel annotation, only 30 images were selected for ground truth preparation in this sub-dataset. Despite its limited size, the Macretina-BV dataset serves as a valuable resource for training and evaluating deep learning models for medical image segmentation in automated ROP severity analysis.

To ensure the dataset covers complex and severe vascular abnormalities, we primarily included retinal images with Aggressive ROP (A-ROP) cases. Since it is characterized by the rapid growth of widened (dilated) and wavy (tortuous) blood vessels, each segmentation mask was precisely validated by ROP experts to accurately capture the growth of those dilated, tortuous, and abnormal blood vessels. The Macretina-BV dataset provides segmentation masks for Aggressive ROP (A-ROP) cases, which are absent in existing ROP blood vessel segmentation datasets, making it a clinically valuable resource for training AI models to handle challenging real-world cases.

3.2.2 Image Pre-Processing

Our proposed architecture presents a system for efficient diagnosis of Retinopathy of Prematurity (ROP) that uses unprocessed, real world data - retinal fundus images of premature babies collected from a local hospital. These real-world images have very low contrast and poor illumination. Also, processing infant's retinal fundus image is more challenging compared to those of adult's as they are noisy and blurred due to the presence of delicate blood vessels. To overcome this challenge, the study uses a set of commonly used pre-processing steps with retinal fundus images, to enhance the image quality as described in figure 3.5.

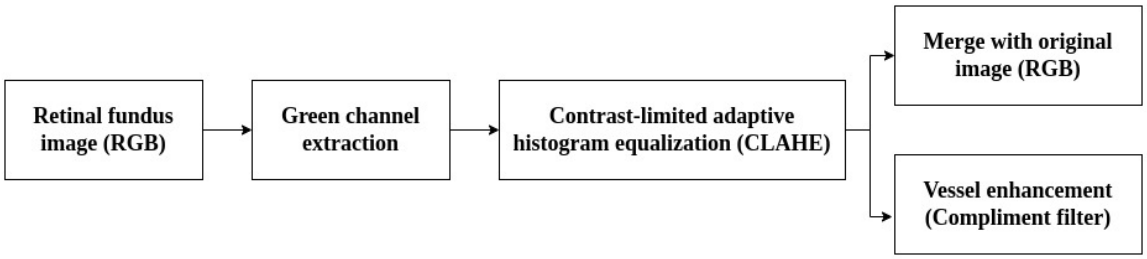


Figure 3.5: Image pre-Processing

It follows feature-specific pre-processing steps to effectively and optimally highlight two morphologically different features ridge/ demarcation line and abnormal blood vessel growth required for Retinopathy of Prematurity (ROP) diagnosis. Ridge detection needs simple pre-processing steps, as the ridge is a relatively well-defined structure present between the vascular and avascular regions. Enhancing contrast and reducing noise is sufficient to detect ridge/demarcation line using our approach. Thus, the study adopts Contrast Limited Adaptive Histogram Equalization (CLAHE) [33] on the green channel of retinal fundus images. The green channel is favored in retinal imaging due to its robustness, it contains better visibility of artifacts while being less prone to noise compared to red and blue channels [34]. By applying CLAHE, we can effectively enhance the contrast of the green channel, thereby improving the visbility of ridge / demarcation line between vascular and avascular regions. While detecting the abnormal growth, dilation and tortuosity in retinal blood vessels is more challenging and requires more advanced pre-processing steps. Once the CLAHE process has

improved the clarity of the green channel, the Compliment Filter [35] is applied to enhance the visualization of blood vessels. The compliment filter is specially designed for the vessel enhancement by selectively focusing on the features of interest while minimizing background noise and improving the contrast of vascular structure.

Combining the outputs from both pre-processing paths allow us to tackle ROP diagnosis more effectively. The CLAHE enhanced green channel helps in ridge detection when merged with original image, while the compliment filter enhances the visibility of blood vessels for the segmentation of blood vessels. These two pathways provide a comprehensive approach to pre-processing retinal fundus images, leading to a more accurate and reliable diagnosis of ROP.

To increase the generalization capacity of the models and reduce overfitting, we applied the basic data augmentation technique to all training sets used in the study. We used horizontal and vertical flips, as they preserve the anatomical structure of the retina while increasing the variability of the dataset. We did not use complex filtering-based augmentation techniques since the dataset consists of retinal fundus images of premature babies, which contains fine and delicate features. Overprocessing such medical images could affect the visibility of clinically relevant features, which are essential for a reliable diagnosis.

3.2.3 Ridge / Demarcation line Detection

Ridge or Demarcation line is the most significant retinal finding to detect ROP from retinal fundus images. Segmentation techniques have been used in the literature to detect the presence of ridge / demarcation line between vascular or avascular region of retina. But the ground truth preparation process required for the segmentation is complex, challenging, and expensive. Direct classification based approach for ridge detection is simple but gives inefficient results due to the lack of localized feature learning, important to detect fine structures like ridge / demarcation line in the retinal fundus images of premature babies. Thus, in our method, we have proposed a hybrid classification-based approach for ridge detection, where we first trained the classifier for 2 class classification with some artificial images, with ridge (similar to the ground truth

used for the ridge segmentation tasks in the literature) and without ridge (complete black image). After this initial training, we fine-tuned the classifier for the original ROP dataset to adapt the actual ridge based features from retinal fundus images.

The initial training with artificial images helps the DCNN-based classifier to effectively learn the different morphological features of the ridge, including its different possible sizes, shapes, and orientations within retinal fundus images. DCNN models follows hierarchical learning, the upper layers of DCNN models learns basic low-level features such as edges, lines, and textures etc, while the deeper layers capture more complex and abstract patterns, including shapes and structures specific to the ridge. This hierarchical learning ability of DCNN models makes them a perfect fit for our two-staged, hybrid classification approach. As in the second stage training process we only train the deeper layers and freeze the upper layers.

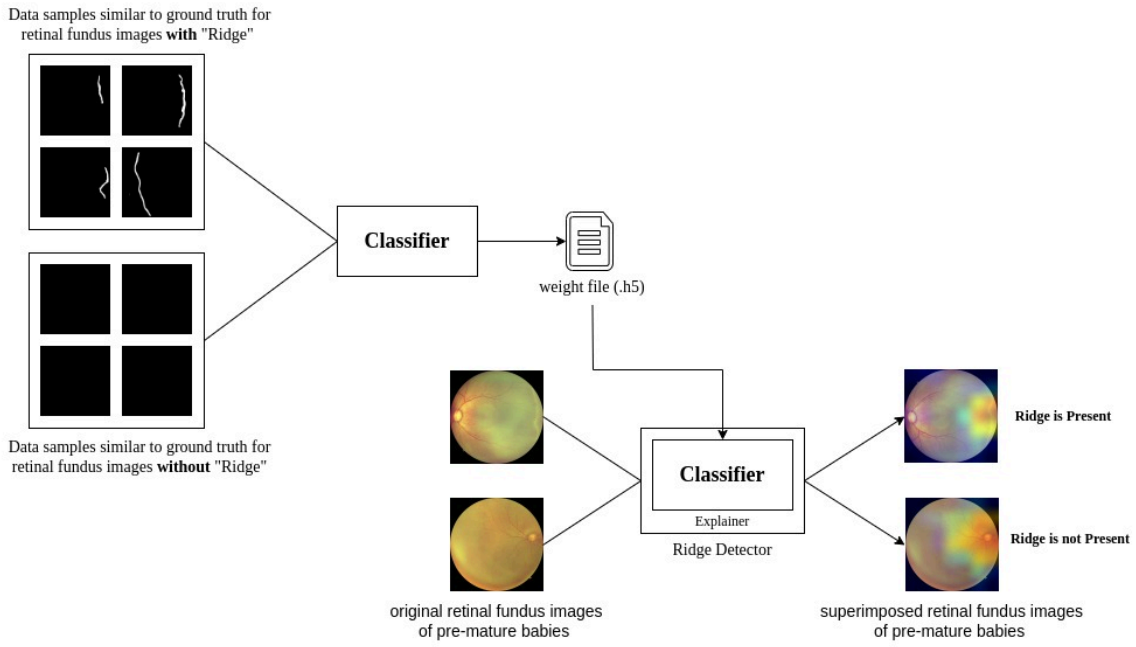


Figure 3.6: Architecture for ridge detection

To support our proposed hybrid classification model we have used Grad-CAM (Gradient-weighted Class Activation Mapping) an Explainable AI (XAI) method to interpret and validate the model's behavior. Grad-CAM generates a heatmap by computing a weighted sum of the feature maps from the last convolutional layer of

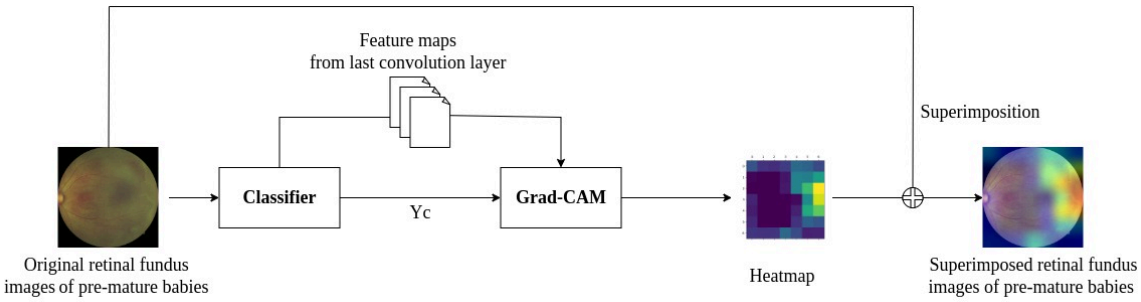


Figure 3.7: Grad-CAM explainations

the classifier, where the weights are the gradients of the respective feature maps with respect to the class score. This Heatmap indicates parts of the input image that are most effective in the model's decision making process. The superimposition of these heatmaps over the original images gives visual explanations of whether the classifier focuses on the ridge or the demarcation line when making decisions.

3.2.4 Blood Vessel Segmentation

Retinal blood vessel map is very important feature in AI-based diagnosis of retinopathy of prematurity (ROP) to accurately find the severity of the disease. Segmentation allows the generation of detailed retinal vascular maps from retinal fundus images, which is very important in order to diagnose the growth of abnormal blood vessels in the retina of premature babies. In recent years, the field of vessel segmentation algorithms for diagnosing retinopathy of prematurity (ROP) has achieved excellent advancements. These advancements have been largely driven by the application of deep convolutional neural networks (DCNNs) [36, 37], which effectively analyze complex patterns of blood vessels in retinal fundus images. Additionally, new techniques such as adversarial training [8] and attention mechanisms [7] have further enhanced the performance of deep learning methods, allowing reliable AI-based ROP diagnosis.

However, deep convolutional neural networks (DCNNs) used for blood vessel segmentation in the literature are typically trained and evaluated on limited datasets, which may not fully capture the variability present in real-world clinical scenarios. To address this, we created a new segmentation dataset called Macretina-BV, which

uniquely includes cases of Aggressive ROP (A-ROP), a severe form of the disease that is often underrepresented or entirely absent in existing studies. By including images from diverse populations and real clinical environments, along with challenging cases like A-ROP, our dataset ensures a more thorough evaluation of segmentation models. The primary objective of this sub-module is to test and validate commonly used DCNN models on our dataset to test their generalization ability across diverse imaging conditions. Retinal blood vessel segmentation in A-ROP cases is challenging due to the presence of highly dilated and tortuous blood vessels, which differ significantly from typical ROP cases. This highly abnormal vascular structure can lead to segmentation inconsistencies, making it difficult for existing methods to accurately create vessel maps in such cases. This approach not only provides critical insights into the robustness and adaptability of existing methods but also evaluates their effectiveness in creating vessel maps for severe form of ROP (i.e AROP).

To perform a comprehensive evaluation of deep convolutional neural network (DCNN) models for retinal blood vessel segmentation, we selected three variants of U-Net architecture such as U-Net, Attention Gate U-Net (AG U-Net), and Residual U-Net (Res U-Net). U-Net architectures have shown effectiveness in medical image segmentation and are frequently used in retinal vessel segmentation tasks. However, their performance on A-ROP (Agressive posterior ROP) cases remains unexplored, which makes our study an important step towards understanding their limitations and potential improvements.

3.2.4.1 U-Net

It follows an encoder-decoder architecture [38], where the encoder (contracting path) progressively reduces spatial dimensions while extracting high-level features, and the decoder (expanding path) restores spatial information using upsampling operations. A key feature of U-Net is its skip connections, which directly transfer the feature maps from the corresponding encoder layers to decoder layers, helping retain spatial information lost during downsampling. Due to its ability to accurately segment complex structures with limited training data, U-Net has been extensively used for

retinal vessel segmentation in AI-based ROP diagnosis.

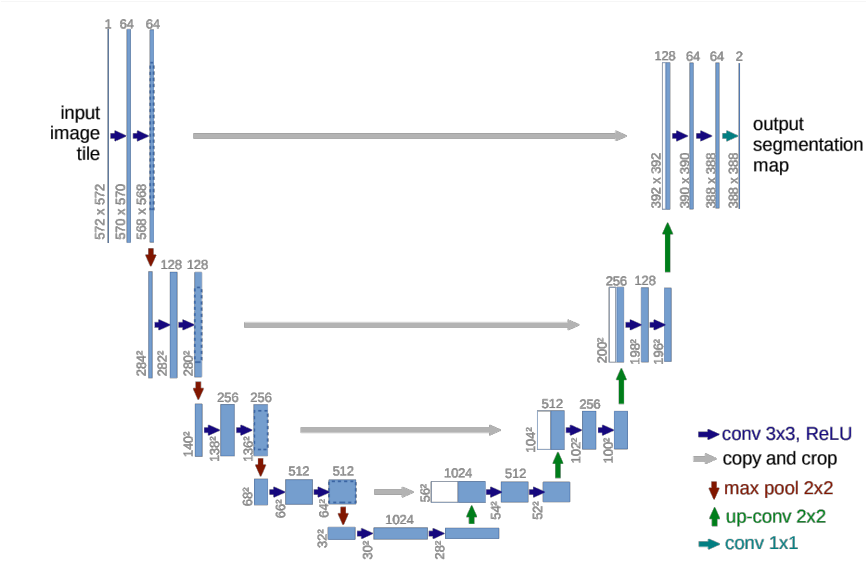


Figure 3.8: U-Net Architecture

3.2.4.2 Attention Gate U-Net (AG U-Net)

It enhances the standard U-Net by incorporating **attention gates** in all four skip connections [39]. This attention mechanism in skip connections helps to ignore the irrelevant features while highlighting important vascular structures, improving segmentation performance in low-contrast and complex regions. This selective feature refinement makes AG U-Net more robust in distinguishing fine blood vessels from its background.

3.2.4.3 Residual U-Net (Res U-Net)

Res U-Net introduces **residual blocks** into all four skip connections, improving gradient flow and feature reuse. By providing **identity mappings**, the residual connections help to solve the vanishing gradient problem, allowing the model to learn deeper features more effectively. This architectural advancement makes Res U-Net well suited for handling the severe vascular abnormalities present in A-ROP cases.

Unlike U-Net and AG U-Net, Res U-Net has not been previously explored for AI-based ROP diagnosis. Additionally, Res U-Net is well suited for training on small

datasets compared to U-Net and AG U-Net. The residual connections improves feature propagation and act like implicit regularizers, which reduces the risk of overfitting. This enables the model to learn richer features even with a limited number of training samples, making it a more robust choice when the large-scale annotated data is not available - like in case of ROP. Furthermore, Res U-Net facilitates faster convergence due to improved gradient flow, leading to more efficient training on medical imaging datasets with limited size.

3.2.5 Zone Separation

The severity of ROP depends on the position of retinal findings in the retina of premature babies. Our method follows the ICROP standards for the zone level classification which is helpful to check the severity of the disease. According to the ICROP standards the retina is divided into three concentric zones centered at center of the optic disc (OD) as shown in figure 3.9, such that the radius of zone-1 is twice the distance between center of optic-disc and maculla, where as the radius of posterior zone-2 is 2 disc diameters beyond zone I radius. Zone-2 area is extending nasally from the outer limit of zone I to the nasal ora serrata and remaining part of retina is reffered as zone-3. To seperate the important retinal zones as defined by the ICROP standards our method follows different steps like OD detection, Computing OD center & OD diameter, Computing OD-to-maculla distance, Computing zone-1 and posterior zone-2 radius, and finally drawing concentric circles with the computed radius for zone-1 and posterior zone-2. However it should be noted that this method only covers clinically severe zones given by ICROP standards and cannot cover the peripheral retinal regions required to identify Zone-2 or above. Therefore, the proposed zone seperator method focuses on detecting the severe ROP cases. The accuracy of zone level ROP classification depends upon how accurately the center and diameter of OD is computed. In our method to compute center and diameter of OD we are using output of OD detection, that is, the coordinates of bounding boxes. Thus to compute OD center and OD diameter more precisely, the accuracy in OD detection is very important.

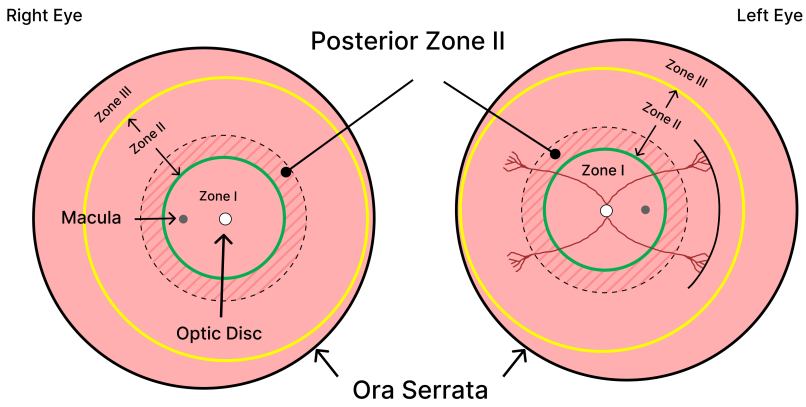


Figure 3.9: ROP zone mapping according to ICROP standards

Optic disc (OD) detection models used in the literature [8] for Retinopathy of Prematurity (ROP) diagnosis are trained on retinal fundus image datasets designed for detecting diseases like diabetic retinopathy, glaucoma, or age-related macular degeneration. Although these models perform well for OD detection in the retinal fundus images of adults, but the differences in eye anatomy between adults and premature babies such as the smaller optic disc size and different retinal proportions can affect their accuracy when applied to neonatal images. This creates a need to adapt or fine-tune these models specifically for preterm infants to make sure that they work effectively for ROP diagnosis. In our proposed method we are using neonatal retinal fundus image datasets especially designed for ROP diagnosis to train, test and validate the OD detection model. In addition to this we have also used some retinal fundus images without OD while training the OD detection model, as in real time ROP diagnosis, retinal fundus images are captured from different angles thus, OD is not visible in all retinal fundus images.

For OD detection in this work we are using YOLOv5 which is highly efficient and accurate, deep learning architecture commonly used for object detection tasks in healthcare. YOLOv5 provides the perfect balance between the speed, accuracy, and efficiency, making it best fit for many realtime and resource constrained applications, such as mobile health (mHealth) and telemedicine. To finetune the YOLOv5 model

for OD detection we have created a large dataset by merging some publicly available datasets related to ROP. Details of all datasets are given in Table 1. After merging these datasets we manually annotated the OD bounding boxes to prepare dataset for OD detection. We also applied some data augmentation techniques such as horizontal flip, vertical flip and rotation to our training data.

The results of OD detection (i.e Bounding box parameters : Xmin, Xmax, Ymin & Ymax) is then used to compute the diameter and center of the OD as described in the figure 3.10. The center of the OD is calculated using equation 1 & 2 and the diameter of the OD is calculated using equations 3,4 & 5.

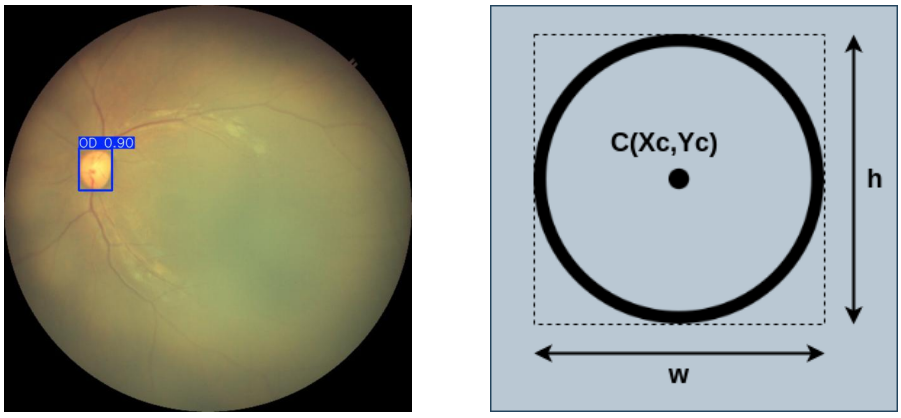
$$X_c = \frac{x_{\min} + x_{\max}}{2} \tag{3.1}$$

$$Y_c = \frac{y_{\min} + y_{\max}}{2} \tag{3.2}$$

$$h = x_{\max} - x_{\min} \tag{3.3}$$

$$w = y_{\max} - y_{\min} \tag{3.4}$$

$$d_{\text{OD}} = \text{avg}(h, w) \tag{3.5}$$



(a) OD detection using YOLOv5

(b) Center and diameter of the OD

Figure 3.10: Illustration of the proposed methodology with Zone Detection and another related image.

The visibility of macula and fovea in retinal fundus images of premature babies

is often compromised due to various factors, including fundus pigmentation and underdevelopment associated with prematurity. Thus, to compute the distance between OD and Fovea (ODF) we have used the clinical findings by [40] which gives the mean value for the ratio between ODF and mean OD diameter (DD) as 3.76, with this ODF/DD ratio and the computed OD diameter we are computing ODF distance for every retinal fundus images using the equation 3.6. where c is pixel to mm constant.

$$\text{ODF} = 3.76 \cdot d_{\text{OD}} \cdot c \tag{3.6}$$

This ODF distance is further used in our method to compute the zone-1 & posterior zone-2 radius according to ICROP standards using equations 3.7 & 3.8 respectively.

$$R_{\text{Zone1}} = 2 \cdot \text{ODF} \tag{3.7}$$

$$R_{\text{Pzone2}} = R_{\text{Zone1}} + 2 \cdot d_{\text{OD}} \tag{3.8}$$

These zone-1 and posterior zone-2 radii are used to draw concentric circles, centered at the center of optic disc to visualize the critical zones for ROP on the retinal funuds images. These concentric circles are drawn using the OpenCV library, as shown in figure, where the blue and green circle covers the zone-1 and posterior zone-2 area, respectively.

Chapter 4

Experimentation and Result

4.0.1 Image Preprocessing

Pre-processing steps used in the proposed approach plays a crucial role in enhancing the quality of retinal images, significantly impacting the diagnosis of Retinopathy of Prematurity (ROP). The results of this study showcase the effectiveness of applying Contrast Limited Adaptive Histogram Eqalization (CLAHE) on the green channel to improve the visibility of ridge, followed by the application of the compliment filter for enhanced vessel segmentation.

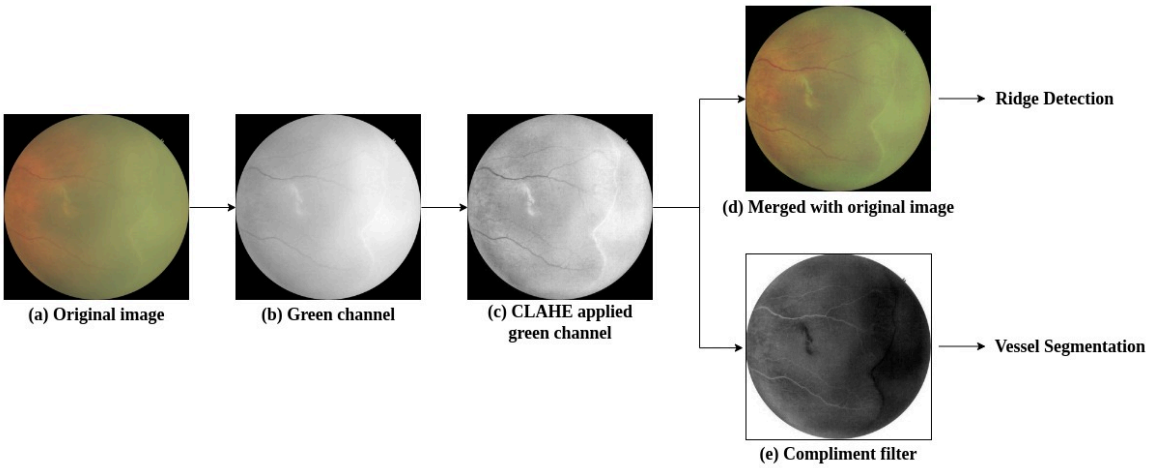


Figure 4.1: Output images of different stages of preprocessing of original color (RGB) image

As shown in the figure 4.1(c), the application of Contrast Limited Adaptive Histogram Equalization (CLAHE with clipLimit of 2.5 & tileGridSize of (8, 8)) on the

green channel significantly enhanced the contrast and visibility of key structures in retinal images. This improvement led to better differentiation between vascular and avascular regions, highlighting the ridge or demarcation line, crucial for diagnosing Retinopathy of Prematurity (ROP). After processing, the enhanced green channel was merged with the original RGB image, resulting in clear visibility of the ridge structure, as presented in figure 4.1(d). This improvement allows Deep Convolutional Neural Networks (DCNNs) to effectively detect the demarcation line with our approach, making it more identifiable than in unprocessed images. Subsequently, the Compliment Filter was applied to further enhance visibility of blood vessels, as shown in figure 4.1(e). This filter efficiently highlighted vascular structures while reducing background noise, leading to improved segmentation of abnormal vessels. Overall, the combined results from CLAHE and the Compliment filter present a comprehensive pre-processing step that effectively addresses the challenges by enhancing critical features for accurate and effective ROP diagnosis.

4.0.2 Ridge / Demarcation line Detection

To validate our hybrid classification-based approach for ridge / demarcation line detection, we fine-tuned five standard convolutional neural networks (CNNs) commonly used in ROP diagnosis: VGG-16, VGG-19 [41], MobileNet [42], MobileNet-V2 [43] and DenseNet201 [44]. These models were trained on existing public ROP datasets [14][45] and then validated on Macretina-ridge dataset for the simple ridge detection task. While DenseNet201 outperformed the other four models in terms of classification metrics with an accuracy of 80.86% over the test dataset, but its results with Grad-CAM - Explainable AI (XAI) technique, were poor. The heat maps generated by this model consistently failed to highlight the critical regions containing the ridge / demarcation line, instead focusing on non-relevant regions of the retina, thus questioning its reliability for real-world applications.

To overcome this problem of a simple classification, we introduce a two-step classification approach which gave significantly better results in terms of both quantitative metrics and interpretability. In this second approach, firstly we trained CNN architec-

tures with some synthetic data (300 - ROP & 300 - NOROP) similar to the masks used in the ridge-segmentation tasks to teach the physical forms and structural characteristics of ridge / demarcation line. Then, we fine-tuned these pre-trained models on the public ROP dataset (of retinal fundus images) for ridge/demarcation line detection task. The test results for this secondary training are presented in Table **??**. With the implementation of our two-step classification approach, significant improvements were observed in the quality of the Grad-CAM heatmaps as shown in Figure 4.2. The revised model generates heatmaps that accurately highlights the ridge/ demarcation lines within the retinal fundus images over superimposition. As shown in the figure , the progressive improvement in the visualization outcomes indicates that initial training with synthetic data has helped the model to gain a deeper understanding of the ridge structures. As a result, the Grad-CAM output not only gives the interpretability but also improves the confidence in the model's decision-making capabilities.

Models	**Accuracy**	**Precision**	**Recall**	**F1 Score**
VGG16	0.8280	0.8036	0.9000	0.8491
InceptionV3	0.7957	0.8163	0.8000	0.8081
DenseNet201	0.9032	0.8868	0.9400	0.9126
MobileNet V1	0.8280	0.9048	0.7600	0.8261
MobileNet V2	0.8602	0.8246	0.9400	0.8785

Table 4.1: Performance of CNN models for Ridge Detection on Macretina-Ridge Dataset

Initial training with synthetic data was done over 20 epochs, whereas the final training was done for 50 epochs. Both these trainings were done with adam optimizer, batch size - 8, learning rate of 0.0001 and softmax - classifier activation function at the output layer. We also added the dropouts of 30-50% in the fully connected layers of CNN models to avoid the problem of overfitting and improve their generalization capacity. The training process lasted approximately 30 minutes, utilizing dual GPU T4 accelerators.

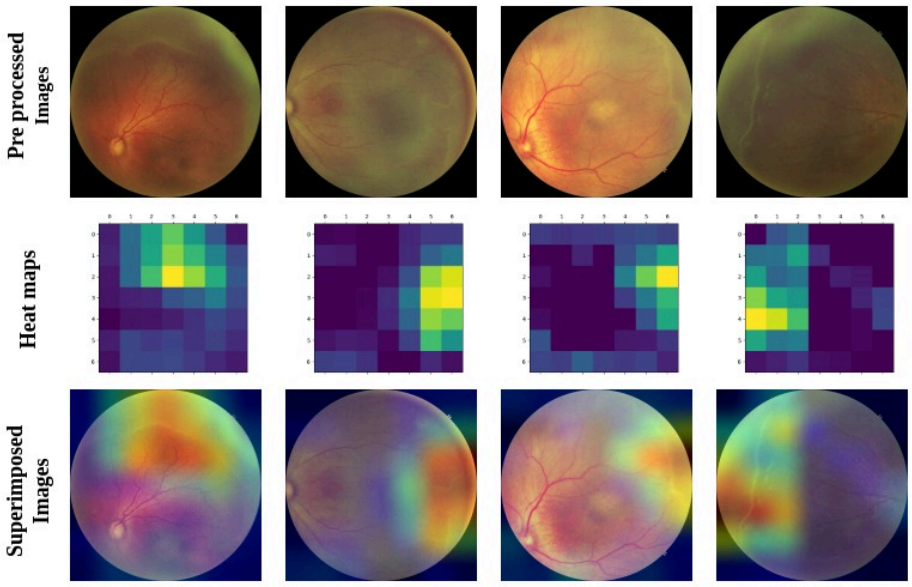


Figure 4.2: Grad CAM interpretation for ridge detection using DenseNet201

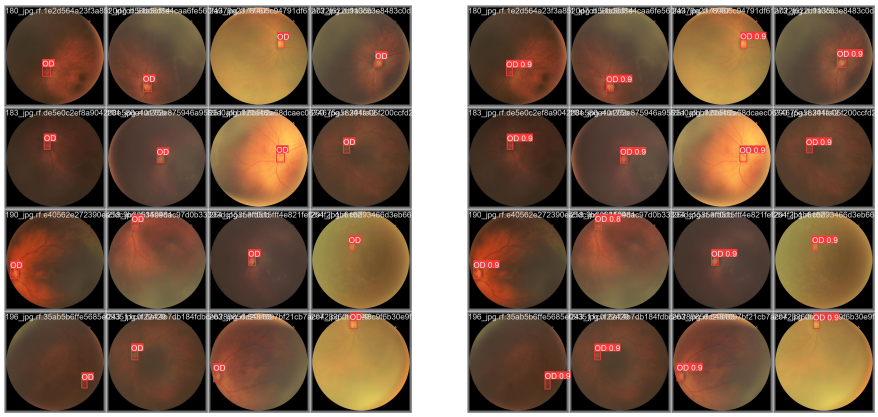
4.0.3 Zone Separation

For optic disc (OD) detection, we used YOLOv5[46], a well-known object detection model known for its speed and accuracy. A recent study on ROP diagnosis [8] presented the effectiveness of YOLOv5 for optic disc detection by highlighting its advantages over YOLOv3 and YOLOv4, particularly in resource-constrained settings such as mHealth and telemedicine. Based on these findings, we chose only YOLOv5 for OD detection on the Macretina-OD dataset.

Model	**mAP@0.5**	**mAP@0.5:0.95**	**Precision**	**Recall**
YOLOv5	0.994	0.800	0.986	0.993

Table 4.2: Performance of YOLOv5 on the Macretina-OD dataset.

The optic disc plays a crucial role in zone-based ROP classification, making its accurate localization essential for automated ROP diagnosis. Zone-based analysis helps to determine how far the abnormal blood vessels have developed from the optic disc. The more posterior the disease (like in Zone I), the more severe and high-risk the ROP is. Therefore, Identifying the correct zone is important for assessing disease progression and urgency of treatment. We used Macretina-OD dataset to evaluate the



(a) Labels (b) Predictions

Figure 4.3: Object Detection Results

performance of YOLOv5 model (Fig. 4.3(b)) for optic disc detection. It consists of 500 images with manually annotated bounding boxes for the optic disc (Fig. 4.3(a)). For training, we fine-tuned the YOLOv5 model — pre-trained on the COCO dataset. The images were resized to 640 × 640 pixels, and the dataset was split into training and testing sets in the ratio 80:20 with a fixed random state of 42. We used a GPU-T4 system for training the model over 50 epochs with a batch size of 16, learning rate of 0.0001, using Adam optimizer. The results of optic disc (OD) detection on the Macretina-OD dataset are presented in Table 4.2. The results demonstrate the acceptability of the dataset in training deep learning models for optic disc detection, a crucial step used in AI-assisted ROP diagnosis.

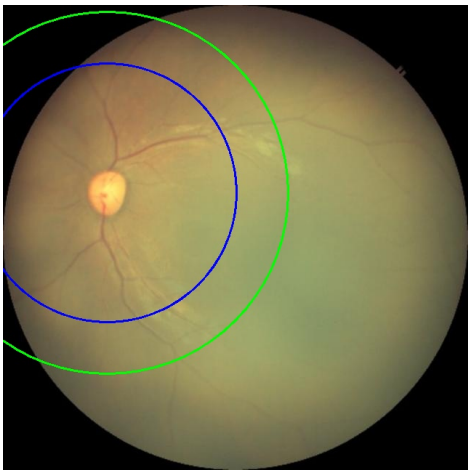


Figure 4.4: Automated mapping of Zone 1 (blue) and Posterior Zone 2 (green)

Successful mapping of severe zones for ROP diagnosis is very important, as it directly informs clinical decision making for treatment options. Thus, following the OD detection, we implemented a zone separation logic that accurately maps medically severe retinal zones according to ICROP standards. The separation method utilizes radii for Zone 1 and posterior Zone 2, computed using the location and diameter of the optic disc within the retinal fundus images (given by Equations 3.7 & 3.8). In the results of our zone separation logic (Figure-4.4), both concentric circles were drawn using OpenCV's library functions function. Zone 1 represented by blue concentric circle centered on the optic disc. This zone is the most critical, as it covers the optic nerve and fovea; any vascular abnormalities in this region indicate a high risk of severe ROP progression and visual impairment. Posterior Zone 2 is represented by a green concentric circle, extending beyond Zone 1 and remaining close to the central retina. This subregion of Zone 2 is particularly important in the clinical evaluation of Aggressive Posterior ROP (AP-ROP). This zone separation logic is validated by senior pediatric ophthalmologists, ensuring that the separation of critical zones is not only precise but also aligned with clinical practices.

4.0.4 Blood Vessel Segmentation

For blood vessel segmentation, we used three DCNN models: U-Net [38], AG (Attention Gate) U-Net [39, 7] and Res U-Net. Accurate blood vessel segmentation is a crucial step in ROP severity analysis, as it helps to analyze vascular abnormalities related with the disease progression. We selected U-Net based models for this experiment because it has been extensively used in the literature for blood vessel segmentation. Due to its encoder-decoder architecture with skip connections, which helps to preserve fine details and spatial information. U-Net based models are well-suited for medical image segmentation tasks where precise boundary detection is very important.

All three models were initially trained on publicly available ROP (HVDROPDB-BV [14], RBV segmentation(ROP) [11]) and Non-ROP (ARIA[47], DRIVE[48], HRF[49]) datasets specifically designed for blood vessel segmentation and then validated on the Macretina-BV dataset. All images were resized to 480 $\times$ 480 pixels.

Model	Accuracy	Dice Coefficient	SSIM
U-Net	0.9624	0.4442	0.9981
AG U-Net	0.9634	0.4571	0.9983
Res U-Net	0.9710	0.5201	0.9987

Table 4.3: Performance of U-Net, AG U-Net and Res U-Net on the Macretina-BV dataset.

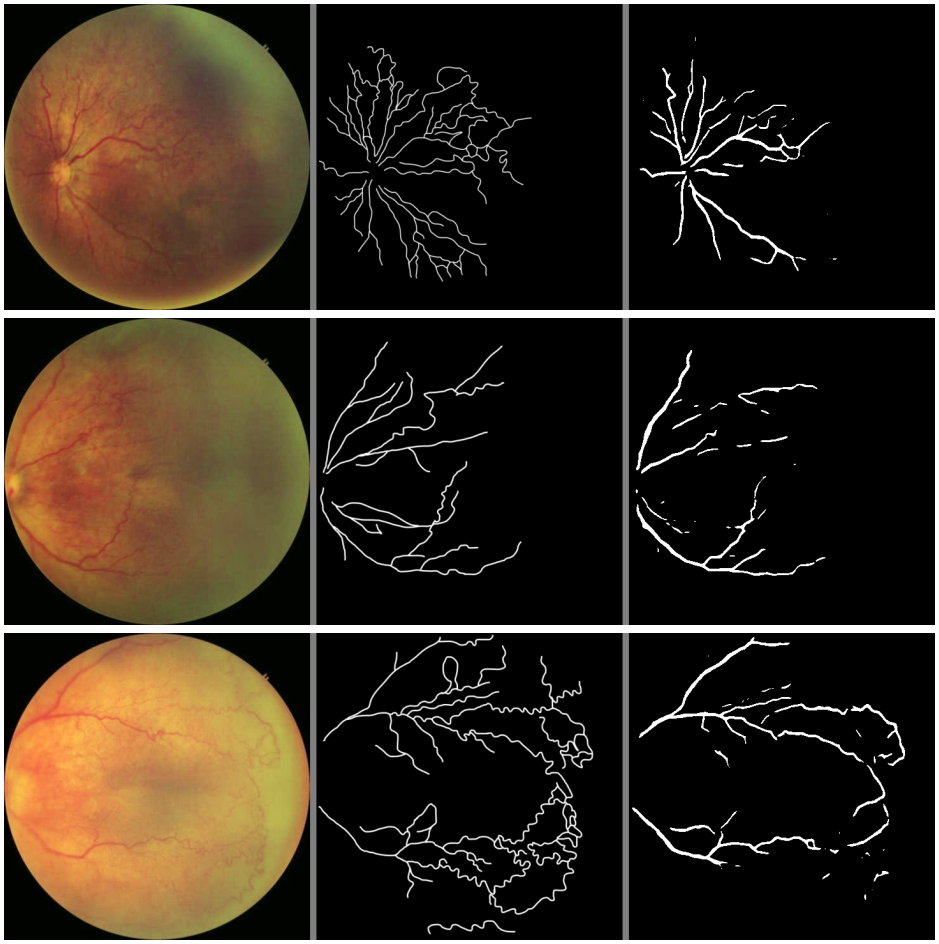


Figure 4.5: Segmentation Results

Training was performed for 100 epochs using the Adam optimizer with a learning rate of 0.0001. We used the Dice Binary Cross-Entropy (DiceBCE) loss function to balance pixel-wise classification and overall segmentation performance, as blood vessels occupy a small portion of the retinal fundus images. The training was carried out on a T4 GPU system, with a batch size of 2, to follow the computational constraints while ensuring effective learning. To measure the model's performance, we used standard segmentation metrics such as Accuracy, SSIM (Structural Similarity Index Measure)

and Dice coefficient - each evaluating how closely the predicted masks align with the ground truth. Sample segmentation outputs for images from validation set, including the original input, ground truth annotation, and the model's predicted mask is shown in Fig. 4.5. The experimental results (presented in the Table 4.3) demonstrate that the Macretina-BV dataset is effective in developing more robust AI-based segmentation models for ROP severity analysis, highlighting its potential for future research and clinical applications.

Chapter 5

Conclusion and Future Work

This chapter details the aspects of extension of the proposed methodology followed by concluding remarks. Subsequent sections encapsulate detailed descriptions of both.

5.1 Future Work

While our proposed solution offers a complete and interpretable solution for ROP diagnosis, several directions remain open for future work.

First, our blood vessel segmentation module currently uses widely used U-Net-based architectures that perform well on standard ROP cases. In this study, we have explored their performance on AROP cases, which are characterized by highly dilated and tortuous vessels. The results show room for improvement on these complex cases. In future, we plan to enhance segmentation accuracy by exploring attention-based models and domain adaptation techniques to capture severe cases like AROP.

Second, although the current ridge detection module leverages synthetic data for initial training, the quantity and diversity of synthetic samples can be further improved using generative models like GANs or diffusion models. This could improve the generalization capacity of classifiers in early-stage ROP detection.

Third, our current zone mapping logic focuses only on clinically severe ROP zones—specifically zone I and posterior zone II—following ICROP guidelines. While this is sufficient for early detection of high-risk cases, future work could extend our

mapping logic to cover all ROP zones, supported by strong medical validation. Additionally, since zone mapping highly depends on accurate optic disc localization, exploring transformer-based techniques or hybrid approaches that integrate clinical factors could further improve the performance, especially in cases with poor illumination or partial occlusion.

Finally, while our current framework uses only retinal fundus images, incorporating multi-modal inputs—relevant clinical metadata (e.g., birth weight, gestational age) along with retinal fundus images can further improve diagnostic accuracy. Future work could explore how integrating these additional metadata enhances the model's ability to detect disease progression and support more comprehensive clinical decision-making.

5.2 Conclusion

This study presents a deep learning-assisted pipeline for automated diagnosis of Retinopathy of Prematurity (ROP), covering three key modules: ridge detection, zone separation, and blood vessel segmentation. Our work not only focuses on creating separate modules for different diagnostic features but also demonstrates how integrating them into a mobile application can improve the overall diagnostic process. For ridge detection, we proposed a novel two-step hybrid classification approach. Our method uses synthetic data in the initial stage to teach the model the structural patterns of ridge formations, followed by fine-tuning on real fundus images. This two step training significantly improved both classification performance and visual interpretability, as validated using Grad-CAM. Among the models evaluated in two-step classification approach, DenseNet201 achieved the highest accuracy, while MobileNet-V2 achieved the best recall, making it particularly suitable for ROP screening task. In the zone seperation module, we fine-tuned YOLOv5 on the expert-annotated Macretina-OD dataset and achieved best detection accuracy. Building on this, we introduced a clinically validated zone separation logic using ICROP guidelines. This method not only outperforms manual zone annotations in consistency but also enables objective

and explainable zone mapping, especially for severe cases like Zone I ROP or AP-ROP. For blood vessel segmentation, we evaluated widely used U-Net-based models on the Macretina-BV dataset. To check their generalizability, we focused specifically on Aggressive ROP (AROP) cases, which are clinically critical due to the presence of highly dilated and tortuous retinal blood vessels. Res U-Net delivered the best results, demonstrating its robustness in capturing complex vascular patterns found in AROP cases. Our solution outperforms existing AI based ROP diagnosis methods in terms of completeness, accuracy, interpretability, and real-time applicability Table-5.1. All modules were finally tested on our newly introduced expert-annotated Macretina dataset, making them clinically reliable. The Macretina dataset has been made publicly available at Figshare for other researchers. Finally, we integrated all three modules into a mobile application, allowing real-time screening in neonatal care units, particularly in low-resource or telemedicine settings.

Work	ROP Datasets	Networks/ Methods	Zone Separation	Ridge Detection	Blood Vessel Map
[8]	AIIMS[8]	YOLO-v5, Pix2Pix GAN	OD Detection Acc: 98.94%; Zone-1 ROP diagnosis	NA	Segmentation Acc: 96.69%, Dice: 0.455 ± 0.073
[7]	HVDROPDB[14]	AG U-Net	NA	Segmentation Acc: 97.00%, Dice: 0.729	Segmentation Acc: 95.00%, Dice: 0.687
[25]	HVDROPDB[14]	Random Forest, AG U-Net	NA	Classification Acc: 80%; F-score(%): 98, 84, 77, 79 for Normal & Stage 1,2,3	NA
[50]	HVDROPDB[14]	U-Net, Circle Hough Transform	Based on OD Segmentation (Dice: 0.844) & Macula, Acc: 98%	NA	Segmentation Acc: 94.00%, Dice: 0.59
[26]	Private Dataset[26]	U-Net1, U-Net3 (Dual Attention), Least-squares Linear Regression	NA	NA	Vessel Segmentation Dice: 0.8116, Sens: 82.73%; OD Segmentation Dice: 0.9346, Sens: 93.95%; Computed vessel tortuosity
Our Work	Macretina [51]	Res U-Net, YOLO-v5, MobileNetV2	OD Detection Acc: 99.50%; Zone-1 and Post Zone-2 ROP diagnosis	Classification Acc: 93.19%; XAI Validation Acc: 88.69%	Segmentation Dice: 0.4703, Acc: 96.11%

Table 5.1: Comparative analysis of other studies on AI-based ROP diagnosis with our proposed approach

Bibliography

[1] H. Blencowe, S. Cousens, D. Chou, M. Oestergaard, L. Say, A.-B. Moller, M. Kinney, J. Lawn, and B. T. S. P. B. A. G. (see acknowledgement for full list), "Born too soon: the global epidemiology of 15 million preterm births," *Reproductive health*, vol. 10, pp. 1–14, 2013.

[2] A. Jana, "Correlates of low birth weight and preterm birth in india," *PLoS One*, vol. 18, no. 8, p. e0287919, 2023.

[3] P. Delhi, "Economic survey highlights thrust on rural development," Press Information Bureau, 2023, available at: https://pib.gov.in/pressreleasepage.aspx?prid=1894901.

[4] U. Tekchandani, D. Katoch, and M. R. Dogra, "Five-year demographic profile of retinopathy of prematurity at a tertiary care institute in north india," *Indian Journal of Ophthalmology*, vol. 69, no. 8, pp. 2127–2131, 2021.

[5] M. F. Chiang, G. E. Quinn, A. R. Fielder, S. R. Ostmo, R. P. Chan, A. Berrocal, G. Binenbaum, M. Blair, J. P. Campbell, A. Capone Jr *et al.*, "International classification of retinopathy of prematurity," *Ophthalmology*, vol. 128, no. 10, pp. e51–e68, 2021.

[6] P. Sen, S. Jain, and P. Bhende, "Stage 5 retinopathy of prematurity: An update," *Taiwan journal of ophthalmology*, vol. 8, no. 4, pp. 205–215, 2018.

[7] R. Agrawal, S. Kulkarni, R. Walambe, M. Deshpande, and K. Kotecha, "Deep dive in retinal fundus image segmentation using deep learning for retinopathy

of prematurity," *Multimedia Tools and Applications*, vol. 81, no. 8, pp. 11 441–11 460, 2022.

[8] V. Kumar, H. Patel, K. Paul, and S. Azad, "Deep learning-assisted retinopathy of prematurity (rop) screening," *ACM Transactions on Computing for Healthcare*, vol. 4, no. 3, pp. 1–32, 2023.

[9] R. Sivakumar, M. Eldho, C. Jiji, A. Vinekar, and R. John, "Diagnosis of plus diseases for the automated screening of retinopathy of prematurity in preterm infants," in *2016 11th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2016, pp. 408–413.

[10] B. A. Scruggs, R. P. Chan, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Artificial intelligence in retinopathy of prematurity diagnosis," *Translational vision science & technology*, vol. 9, no. 2, pp. 5–5, 2020.

[11] J. Ji, "Retinal blood vessel segmentation (rop)," 2023. [Online]. Available: https://doi.org/10.17632/prcy36j53v.2

[12] J. Wang, J. Ji, M. Zhang, J.-W. Lin, G. Zhang, W. Gong, L.-P. Cen, Y. Lu, X. Huang, D. Huang *et al.*, "Automated explainable multidimensional deep learning platform of retinal images for retinopathy of prematurity screening," *JAMA network open*, vol. 4, no. 5, pp. e218 758–e218 758, 2021.

[13] X. Zhao, S. Chen, S. Zhang, Y. Liu, Y. Hu, D. Yuan, L. Xie, X. Luo, M. Zheng, R. Tian *et al.*, "A fundus image dataset for intelligent retinopathy of prematurity system," *Scientific Data*, vol. 11, no. 1, p. 543, 2024.

[14] R. Agrawal, R. Walambe, K. Kotecha, A. Gaikwad, C. M. Deshpande, and S. Kulkarni, "Hvdropdb datasets for research in retinopathy of prematurity," *Data in Brief*, vol. 52, p. 109839, 2024.

[15] S. N. Patel, T. Al-Khaled, K. B. Kang, K. E. Jonas, S. Ostmo, C. V. Ventura, M. A. Martinez-Castellanos, R. G. A. S. Anzures, J. P. Campbell, M. F. Chiang

et al., "Characterization of errors in retinopathy of prematurity diagnosis by ophthalmologists-in-training in middle-income countries," *Journal of Pediatric Ophthalmology & Strabismus*, vol. 60, no. 5, pp. 344–352, 2023.

[16] K. Sabri, "Global challenges in retinopathy of prematurity screening: modern solutions for modern times," *Pediatrics*, vol. 137, no. 1, 2016.

[17] Z. Tan, M. Isaacs, Z. Zhu, S. Simkin, M. He, and S. Dai, "Retinopathy of prematurity screening: A narrative review of current programs, teleophthalmology, and diagnostic support systems," *Saudi Journal of Ophthalmology*, vol. 36, no. 3, pp. 283–295, 2022.

[18] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Computerized medical imaging and graphics*, vol. 31, no. 4-5, pp. 198–211, 2007.

[19] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps: Automation of decision making*, pp. 323–350, 2017.

[20] J. E. Reid and E. Eaton, "Artificial intelligence for pediatric ophthalmology," *Current opinion in ophthalmology*, vol. 30, no. 5, pp. 337–346, 2019.

[21] R. Sayres, N. Hammel, and Y. Liu, "Artificial intelligence, machine learning and deep learning for eye care specialists," *Annals of Eye Science*, vol. 5, pp. 18–18, 2020.

[22] O. Attallah, "Gabrop: Gabor wavelets-based cad for retinopathy of prematurity diagnosis via convolutional neural networks," *Diagnostics*, vol. 13, no. 2, p. 171, 2023.

[23] A. Bai, C. Carty, and S. Dai, "Performance of deep-learning artificial intelligence algorithms in detecting retinopathy of prematurity: A systematic review," *Saudi Journal of Ophthalmology*, vol. 36, no. 3, pp. 296–307, 2022.

[24] P. Maitra, P. K. Shah, P. J. Campbell, and P. Rishi, "The scope of artificial intelligence in retinopathy of prematurity (rop) management," *Indian Journal of Ophthalmology*, vol. 72, no. 7, pp. 931–934, 2024.

[25] R. Agrawal, S. Kulkarni, A. Gaikawad, C. M. Deshpande, K. Kotecha, and R. Walambe, "Deep learning framework for classification and explanation of early rop stages from hvdropdb dataset," in *2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)*. IEEE, 2024, pp. 1–6.

[26] Y. Wang, X. Zheng, C. He, J. Zhang, and S. Huang, "Improved segmentation of infant retinal images and quantitative vascular analysis," *IEEE Access*, 2024.

[27] O. Attallah, "Diarop: automated deep learning-based diagnostic tool for retinopathy of prematurity," *Diagnostics*, vol. 11, no. 11, p. 2034, 2021.

[28] J. Almeida, J. Kubicek, M. Penhaker, M. Cerny, M. Augustynek, A. Varysova, A. Bansal, and J. Timkovic, "Enhancing rop plus form diagnosis: An automatic blood vessel segmentation approach for newborn fundus images," *Results in Engineering*, vol. 24, p. 103054, 2024.

[29] C. Liu, X. Yin, D. Huang, Y. Xu, S. Li, C. Yu, Y. Zhang, and Y. Deng, "An interpretable prediction model for the risk of retinopathy of prematurity development based on machine learning and shapley additive explanations (shap)," 2023.

[30] J. Wang, J. Ji, M. Zhang, J.-W. Lin, G. Zhang, W. Gong, L.-P. Cen, Y. Lu, X. Huang, D. Huang, T. Li, T. K. Ng, and C. P. Pang, "Automated explainable multidimensional deep learning platform of retinal images for retinopathy of prematurity screening," *JAMA Network Open*, vol. 4, no. 5, pp. e218 758–e218 758, 05 2021.

[31] Forus Health. (2016) 3nethra neo - retinal imaging system for neonatal eye screening. [Online]. Available: https://forushealth.com/3nethra-neo/

[32] P. K. Shah, V. Narendran, and N. Kalpana, "Aggressive posterior retinopathy of prematurity in large preterm babies in south india," *Archives of Disease in Childhood-Fetal and Neonatal Edition*, vol. 97, no. 5, pp. F371–F375, 2012.

[33] K. Zuiderveld, *Contrast limited adaptive histogram equalization.* USA: Academic Press Professional, Inc., 1994, p. 474–485.

[34] M. J. Alwazzan, M. A. Ismael, and A. N. Ahmed, "A hybrid algorithm to enhance colour retinal fundus images using a wiener filter and clahe," *Journal of Digital Imaging*, vol. 34, no. 3, pp. 750–759, 2021.

[35] S. Dash, S. Verma, Kavita, M. S. Khan, M. Wozniak, J. Shafi, and M. F. Ijaz, "A hybrid method to enhance thick and thin vessels for blood vessel segmentation," *Diagnostics*, vol. 11, no. 11, p. 2017, 2021.

[36] B. Goutam, M. F. Hashmi, Z. W. Geem, and N. D. Bokde, "A comprehensive review of deep learning strategies in retinal disease diagnosis using fundus images," *IEEE Access*, vol. 10, pp. 57 796–57 823, 2022.

[37] A. Jafarizadeh, S. Maleki, P. Pouya, N. Sobhi, M. Abdollahi, S. Pedrammehr, C. Lim, H. Asadi, R. Alizadehsani, R. Tan *et al.*, "Current and future roles of artificial intelligence in retinopathy of prematurity. arxiv 2024," *arXiv preprint arXiv:2402.09975*, 2024.

[38] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18.* Springer, 2015, pp. 234–241.

[39] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[40] D. J. De Silva, K. D. Cocker, G. Lau, S. T. Clay, A. R. Fielder, and M. J. Moseley, "Optic disk size and optic disk-to-fovea distance in preterm and full-term infants," *Investigative ophthalmology & visual science*, vol. 47, no. 11, pp. 4683–4686, 2006.

[41] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[42] A. G. Howard, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[43] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[44] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[45] J. Timkovič, J. Nowaková, J. Kubíček, M. Hasal, A. Varyšová, L. Kolarčík, K. Maršolková, M. Augustynek, and V. Snášel, "Retinal image dataset of infants and retinopathy of prematurity," *Scientific Data*, vol. 11, no. 1, p. 814, 2024.

[46] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, K. Michael, J. Fang, C. Wong, Z. Yifu, D. Montes *et al.*, "ultralytics/yolov5: v6. 2-yolov5 classification models, apple m1, reproducibility, clearml and deci. ai integrations," *Zenodo*, 2022.

[47] P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel detection and measurement using wavelets and edge location refinement," *PloS one*, vol. 7, no. 3, p. e32435, 2012.

[48] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[49] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *International journal of biomedical imaging*, vol. 2013, no. 1, p. 154860, 2013.

[50] R. Agrawal, S. Kulkarni, R. Walambe, and K. Kotecha, "Assistive framework for automatic detection of all the zones in retinopathy of prematurity using deep learning," *Journal of Digital Imaging*, vol. 34, no. 4, pp. 932–947, 2021.

[51] U. Trivedi, A. Srivastava, and P. Mahajan, "Macretina: a dataset, to support deep learning assisted retinopathy of prematurity diagnosis," *Scientific Reports*, 2025.