A comprehensive study for predicting eyesight disease using ML

by

Tanvir Islam Sayem
20301360
Fouzia Rahman Sara
20101122
Poroma Biswas
20201084
Debabrata Bhowmick
20301374

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
2024

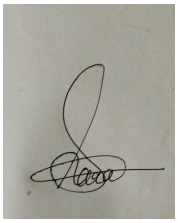
Declaration

It is hereby declared that

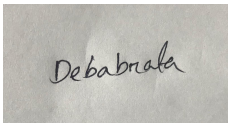
1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

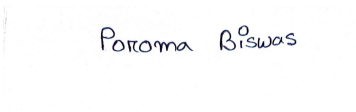
Tanvir Islam Sayem
20301360



Fouzia Rahman Sara
20101122



debabrata Bhowmick
20301374



Poroma Biswas
20201084

Approval

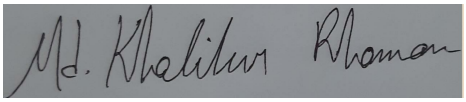
The thesis/project titled "A comprehensive study for predicting eyesight disease using ML" submitted by

1. Tanvir Islam Sayem (20301360)
2. Fouzia Rahman Sara (20101122)
3. poroma Biswas (20201084)
4. Debabrata Bhowmick (20301374)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on 2024

Examining Committee:

Supervisor:
(Member)



Dr.Khalilur Rhaman,PhD
Professor
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement (Optional)

This is optional, if you don't have an ethics statement then omit this page

Abstract

From mild to severe distant vision impairment caused by untreated conditions including cataract, glaucoma, retinal disease, and diabetic retinopathy, more than 60% of the world's population—exceeding 4.5 billion individuals—requires corrective lenses or treatments for visual and retinal disorders. The fundamental goal of the current study is to create an advanced deep learning (DL) system capable of categorizing retinal pictures into five groups. A deep convolutional neural network (CNN) was used to classify normal eyes, cataracts, glaucoma, retinal illness, and diabetic retinopathy. The dataset, obtained from Kaggle, had 2827 pictures that were randomly divided into training, validation, and testing groups. The TensorFlow object identification framework was used to create many CNN meta-architectures, including YOLOv5, YOLOv7, and InceptionResNet50. The YOLOv5 model showed great development. The YOLOv5 model demonstrated significant progress in detecting the mentioned eye diseases and achieving 0.951 mAP for 7357 images.

Keywords: Eye Diseases ; Machine Learning; Deep Learning; YOLOv7; Prediction; Inception-Resnet50; ; YOLOv5

Dedication (Optional)

Acknowledgement

First and foremost, all praises and thanks are due to Allah, the Most Gracious and the Most Merciful, for His boundless blessings and guidance throughout the completion of this thesis. Without His divine support, none of this would have been possible. We would like to express our sincere gratitude to our supervisor, Dr.Khalilur Rhaman, for his invaluable guidance, unwavering support, and insightful feedback. His expertise in the field and dedication to our academic growth have been instrumental in shaping this research.Dr.Khalilur Rhaman's patience, encouragement, and willingness to devote his time and energy to our progress have truly been inspiring. We would also like to extend our heartfelt appreciation to our co-supervisor, Avijit Biswas for his cooperation and assistance throughout this journey and we highly grateful to Sayantan Roy Sir for Their expertise, constructive criticism, and willingness to share their knowledge have greatly enriched our work. Their guidance and continuous encouragement have played a pivotal role in shaping our research direction and refining our ideas. We are deeply grateful to the entire faculty of our department for their teachings, mentorship, and support. Their collective efforts have nurtured our intellectual growth and contributed to the development of our research skills. Furthermore, we would like to acknowledge the contributions of our friends and colleagues, whose discussions, suggestions, and feedback have been invaluable in shaping our ideas and enhancing the quality of this thesis.Their support and camaraderie have made this journey all the more enjoyable and rewarding. In conclusion, we recognize and acknowledge the support and contributions of all those who have played a part, directly or indirectly, in the completion of this thesis. Their guidance, encouragement, and assistance have been indispensable in our academic and personal growth. We are humbled and grateful for their presence in our lives. May Allah bless each and every one of them with success, happiness, and prosperity in all their future endeavors.

Table of Contents

List of Figures

List of Tables

Nomenclature

Chapter 1

Introduction

Eye diseases represent a significant global health concern, affecting millions of individuals worldwide with various ocular conditions. Conditions like Cataract, Glaucoma, Retinal disease, and diabetic retinopathy contribute to the overall burden of eye-related issues. This research aims to develop a versatile solution by utilizing retinal images to detect various eye disorders, including cataracts, glaucoma, and retinal diseases. Deep convolutional neural networks will be employed for classification. The primary objective is to create a flexible method for identifying eye illnesses, encompassing conditions like cataract, diabetic retinopathy, glaucoma, and retinal diseases through the evaluation of retinal images using convolutional neural networks. This method holds promise for early detection and efficient treatment of these diseases. Figure 1.1 illustrates a retinal identification pattern of a normal eye.

Research Problem

In 2019, the World Health Organization (WHO) reported that over 2.2 billion people globally suffer from visual impairment or blindness. The greatest challenge for visually impaired individuals is navigating their surroundings, whether indoors or outdoors, leading to risky and perilous lives. Early treatment options are available, offering a chance to cure or slow down the progression of diseases, providing patients with additional years of vision. However, in countries like India and Bangladesh, despite the presence of numerous hospitals and eye clinics in urban areas, there is a low doctor-patient ratio, particularly in rural regions where both infrastructure and access to ophthalmologists are lacking. The scarcity of qualified personnel in rural areas poses challenges for implementing community aid programs and conducting eye examinations. The automation of the eye diagnostic system through technological advancements and image analysis can offer a solution. This automation allows for remote image capture and diagnosis, reducing costs and infrastructure requirements. Additionally, it enables doctors to refer patients for more comprehensive guidance. In order to optimize the comprehensive analysis of image color, it is necessary to account for adjustments in the image, including factors such as light scattering, pathology, variability in fundus images, and reflections from the fundus. This includes variations in thickness and framing based on usage, as well as differences in cameras, lighting conditions, shooting angles, and retinal pigmentation. Alternative methods, such as grayscale global normalization, histogram equalization, and histogram specification, are also accessible [5]. The standardization of nearly all

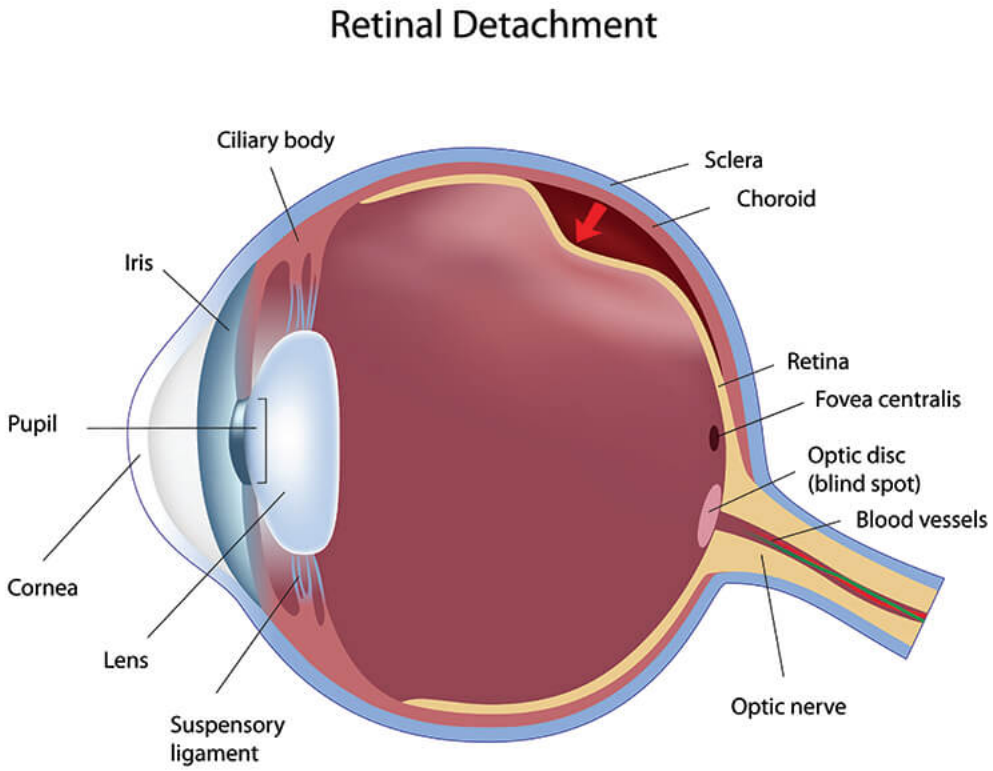


Figure 1.1: Labeled retinal image

color properties is an emerging trend [6].Scientists have created clinical decision support systems to aid in the diagnosis of diabetic retinopathy. Others aid in the diagnosis of age-related macular disorders. These systems take advantage of recent advancements in the processing of digital images and machine learning.Many of these systems take on the role of human specialists. They do, however, specialize in the diagnosis of a specific retinal condition.

Numerous models leverage retinal images to identify, extract, and analyze disease-specific characteristics. This approach requires a deep understanding of the respective conditions and significant effort in constructing effective classifier functions.

CNNs have shown to be incredibly effective as core deep learning approaches for diagnosing eye disorders. This study extends the use of deep CNNs to the classification of eye diseases. The study uses deep learning architectures such as YOLOv5, YOLOv7, and Inception-ResNet50 to extract deep features from a dataset of genuine eye disease photos. The results show that the YOLOv5 algorithm beat the YOLOv7 and Inception-ResNet50 algorithms in the eye disease classification task. Deep convolutional neural networks (CNNs) have made significant strides in the realms of deep learning and visual object recognition. Their ability to extract features without depending on segmented approaches is a notable advantage, streamlining their practical implementation.

To recap, previous investigations on eye disease used typical image processing approaches for recognition, yielding precise results with excellent disease accuracy.

Nonetheless, there are significant downsides and limits, such as:

- The approaches and procedures used in research are sophisticated, highly individualized, and labor and time-intensive.
- There is a great deal of emphasis on spot segmentation.
- The method relies significantly on the extraction of fake features.
- Evaluating the performance of illness recognition is difficult

Research Objectives

Eye diseases are a significant global health concern. They affect millions of individuals and pose challenges to healthcare systems worldwide. Eye diseases are common. They can cause permanent vision loss. This shows the need for more research in this area. Recent advancements in artificial intelligence and computer vision create a growing opportunity. We can now use these technologies to predict and detect eye diseases early. This research aims to address critical gaps in current approaches. It will develop and put in place predictive models. The research will focus on utilizing state-of-the-art deep learning techniques such follow:

- Developing an Efficient YOLOv5 Model for Eye Disease Detection: Create a customized YOLOv5 model optimized for the accurate detection and classification of eye diseases in fundus images
- Curating and Augmenting a Diverse Dataset: Assemble a representative dataset of fundus images, spanning various eye diseases, and enhance model robustness through data augmentation
- Evaluating Model Performance: - Assess the precision, recall, and F1 score of the YOLOv5 model, considering its ability to handle diverse image complexities and disease severities.
- Comparative Analysis with Other Models:Contrast the YOLOv5 model's performance with other cutting-edge deep learning models commonly used in medical image analysis for eye diseases
- Enhancing Model Interpretability: Improve the interpretability of YOLOv5 predictions for eye diseases and explore techniques for providing transparent decision-making visualizations
- Robustness to Image Variability:Investigate the feasibility and effectiveness of deploying the YOLOv5 model in real-world clinical scenarios, addressing challenges in diverse patient populations
- Ethical Considerations and Bias Mitigation:Examine ethical concerns related to using AI models for medical diagnosis, focusing on patient privacy and implementing strategies to reduce biases.

- Integration into User-Friendly Interfaces:Explore integrating the YOLOv5 model into user-friendly interfaces for healthcare professionals, evaluating usability and acceptance in clinical settings
- Proposing Future Enhancements and Research Directions:Suggest potential improvements to the YOLOv5 model based on research findings and provide recommendations for future exploration in eye disease prediction using deep learning.

Chapter 2

Literature Review

Common eye ailments prevalent among the elderly consist of cataract, glaucoma, retinal disease and diabetic retinopathy. An intelligent computer-based system was devised to categorize these eye disorders, employing three distinct types of machine learning algorithms [8]In this part of the document, an overview of recent significant literature is provided. The writers detailed an innovative method for capturing primary features from color retinal images as outlined in reference [7].To locate the optic disk, detect its shape, and offer a more comprehensive feature description in retinal images, Principal Component Analysis (PCA), a modified functional structure approach, and an enhanced fundus coordinate structure are employed. The success rates for these three tasks are 99%, 94%, and 100%, respectively.In an alternate study, scientists employed deep learning through a deep-belief network to identify early-stage glaucoma [9].

Another method involved diagnosing glaucoma by examining the identified optic disc, utilizing wavelet-based statistical and textural features. The algorithm's effectiveness was assessed on two publicly available datasets with different resolutions, showing an accuracy of 96.7% [10]. Treder et al. introduced a model for detecting AMD, utilizing an open-source deep convolutional neural network, and achieving an accuracy of 98% [11]

The authors proposed using a back propagation neural network (BPNN) classification to automatically diagnose cataracts by identifying retinal images in their paper [12]. The BPNN approach assigns images to one of four categories: normal, mild, medium, or severe cataract. Another study [13] did a comprehensive literature review on extracting anatomical information from retinal pictures to aid in the early detection of glaucoma. The scientists assessed existing automatic extraction methods, taking into account variables such as Retinal Nerve Fiber Layer (RNFL) , Neuroretinal Rim Notching, Optic Cup to cup-to-disc ratio (CDR) and others . The study systematically examined chosen studies from 2014 to 2019, omitting those that used the segmented optic disk approach, with the conclusions described in paper [14].

Leveraging the YOLO v5 network [27] introduced a novel deep model designed for object recognition. This advanced method excels in identifying objects of various sizes, ranging from very large to small. The authors specifically recommend employing multiscale models to learn intricate features, considering representations of different dimensions and selecting the most suitable scales for object recognition.Significantly, the suggested model utilizes fewer parameters during the training

phase when juxtaposed with the original YOLO-v5 model architecture. This reduction in parameters contributes to a noteworthy enhancement in accuracy, a conclusion supported by the results obtained from various test scenarios. The proposed model exhibited a slight reduction in YOLO v5-S profiles, decreasing from 7.28 million to 7.26 million parameters. The system detected objects at 30 frames per second. This is slower than YOLO v5 L/X profiles. However, the system detected small vehicles 33% better than the YOLO v5-X profile.
In paper [27] presents an overview of retinal digital analysis concepts, methodologies, and applications. Advancements in image analysis and pattern recognition have enabled the automatic identification of ocular disorders. Examples include glaucoma , age-related macular degeneration and diabetic retinopathy. The study of the link between retinal microvasculature and systemic cardiovascular disorders. The study discusses the fundamentals of retinal digital image analysis. It also covers current strategies for automated landmark segmentation. It discusses advancements in recognizing diabetic retinopathy features. The authors examine the challenges of quantitative measurements from retinal pictures. They discuss the arteriovenous ratio (AVR). They end by suggesting potential uses of fundal image analysis in telemedicine.The objective of this paper [15] was to evaluate the effectiveness of AI-based automated software in diagnosing sight-threatening diabetic retinal degeneration (STDR) and diabetic retinal degeneration (DR) using fundus images captured with a smartphone device.A total of 296 retinal pictures were evaluated, demonstrating that the program had a degree of sensitivity of 95.8% and a preciseness of 80.4%. The study's [16] goal was to diagnose glaucoma using basic elements of analytics (PCA) and a Bayesian classifier. The success rates for optic disk localization were 95.3% for the healthy set and 89.9% for the glaucoma set, with a total accuracy of 78% tested on more than 300 retina images. Gardner et al. [17] created a back-propagation neural network to identify diabetic features in retinal fundus images, aiding in the diagnosis of diabetic retinopathy. The examined retinal features encompassed blood vessels, exudates, and hemorrhages. To enhance accuracy and precision on the training data, we optimized the initial filtering and network parameters. The training dataset comprised 147 diabetic fundus images and 32 normal retina images. Subsequently, the network was assessed using a test dataset containing 200 diabetic and 101 normal photos. In comparison to evaluations by a trained ophthalmologist, the neural network demonstrated a sensitivity of 88.4% and specificity of 83.5% in identifying diabetic retinopathy.The proposed approach stands out by incorporating deep learning, diverging from the approaches documented in the literature. It leverages advanced features, steering clear of commonplace techniques like SVM, K-NN, FCM, or optimization methods. The selection of CNN is strategic for efficiently handling diabetic retinopathy imaging data, given its compatibility with image data. The method introduces novel features and advantages, featuring a straightforward design structure that integrates simple image processing techniques during the preprocessing phase. In contrast, other studies strive for favorable outcomes through the amalgamation and adaptation of diverse techniques. The proposed method in this study utilizes histogram equalization (HE) and the contrast limited adaptive histogram equalization (CLAHE), recognized as effective preprocessing techniques for priming data for CNNs. This study puts forth a targeted and efficient solution with the objective of enhancing the efficiency of classifying retinal fundus images, employing mixed methods that entail high processing requirements.

Chapter 3

Workflow

Deep Convolutional Neural Networks (CNNs) are a subset of neural networks. They have excelled in a variety of computer vision and image processing events. CNNs are utilized in a variety of applications.video processing , Image classification, segmentation, natural language processing ,object recognition and speech recognition are some examples. Their excellent learning ability stems from the fact that they use many feature extraction processes. These processes enable data representation for deep learning technology. The significant growth in data availability and developments in hardware technology have sparked interest in CNN research. This resulted in the birth of novel structures. Experiments with various activation and loss functions are part of this research. Parameter optimization, regularization, and architectural enhancements are also included.

On the flip side, the deep CNN demonstrates a considerable advancement in its capability to represent information. This is accomplished through architectural improvements.Notably, there has been considerable attention given to key concepts such as Feature-Map, spatial explorations ,width based Multi-Connection, depth based CNN, Multi-path based CNN ,Channel explorations and attention based in recent model of deep Convolutional Neural Network (CNN) designs which is depicted visually in Figure 3.1 . The study's goal is to develop a compact deep-learning framework for early identification of eye disorders. Noise and dark areas in fundus imaging may necessitate image pre-processing to improve image quality. Improved image quality improves the model that is suggested for recognizing significant indications of eye disorders. Furthermore, a specific feature extraction technique gives a collection of features, which enables the suggested model to generate results with little computational resources. Image pre-processing procedures, extraction of features and selection, and recognition methods are all part of this.
To improve its object identification performance and representational capacity, YOLOv5 employs a deep convolutional neural network (CNN). YOLOv5's deep design enables it to recognize a variety of patterns in data, abstract complicated information, and learn a hierarchy of characteristics. Applying non-linear mappings and efficiently transforming input information, a deep CNN enables YOLOv5 to significantly enhance its object recognition and representation capabilities across a wide range of sizes, orientations, and circumstances. A wider receptive field is created by the

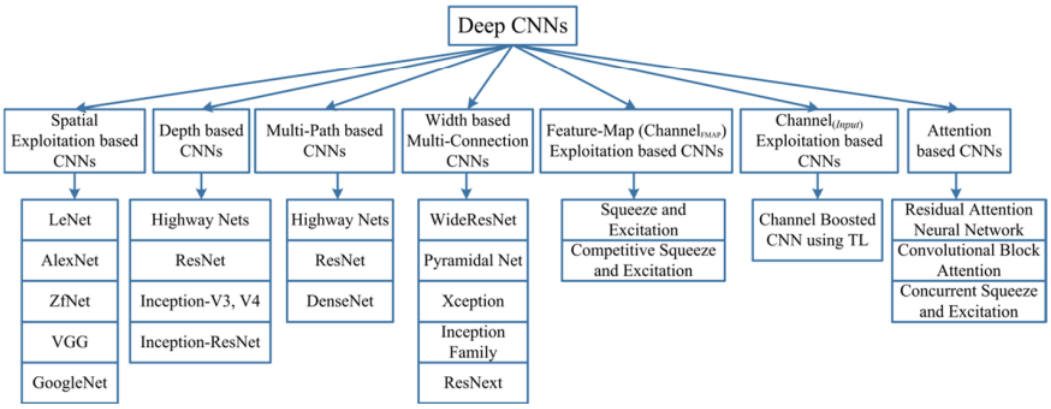


Figure 3.1: Sven Distinct Categories of Deep Convolutional Neural Network

deeper penetration, which improves contextual comprehension. Overall, the incorporation of deep CNN techniques in YOLOv5 enables it to excel in object identification tasks by exploiting advanced architectural aspects, attention mechanisms, and efficient design decisions for both training and inference.
The model excels in image classification, delivering impressive results with minimal computer resource usage. YOLOv5's architecture prioritizes efficient computation, making it well-suited for distributed processing or deployment on hardware with multiple processing units. This adaptability allows the proposed eye diseases detection model to be deployed effectively on mobile devices and other low-powered endpoints. The classification of fundus images in the proposed model is carried out using the YOLOv5 approach, alongside InceptionV3, ResNet50, and YOLOv7, as illustrated in Figure 3.2 .

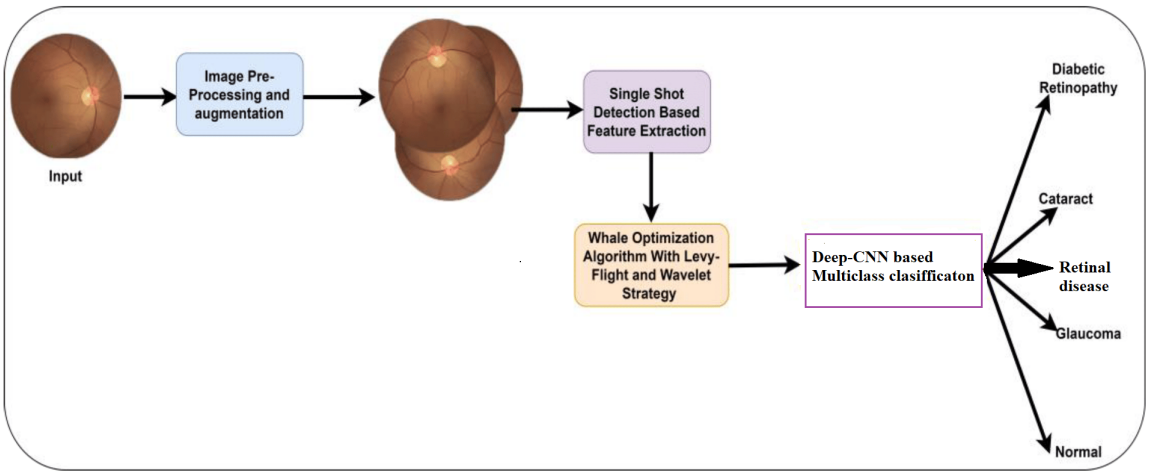


Figure 3.2: Simple workflow of Eye Diseases Detection Model

The model excels in image classification, delivering impressive results with minimal computer resource usage. YOLOv5's architecture prioritizes efficient computation, making it well-suited for distributed processing or deployment on hardware with multiple processing units. This adaptability allows the proposed eye diseases detection model to be deployed effectively on mobile devices and other low-powered endpoints. The classification of fundus images in the proposed model is carried out

using the YOLOv5 approach, alongside InceptionV3, ResNet50, and YOLOv7, as illustrated in Figure 3.2.

Likewise, the trend of incorporating a block of layers as a fundamental structural unit has gained significant popularity. This study not only imparts a fundamental understanding of the components of Convolutional Neural Networks (CNNs) but also delves into the current challenges and applications associated with these advanced CNN architectures

Chapter 4

Description of the data

4.1 Dataset Overview

The dataset is based on primary source and secondary source (obtained from Kaggle [18]).we collect the primary data form the National Institute of Ophthalmology Dhaka,Bangladesh. consists of retinal pictures classified into five categories: Normal, Cataract, Glaucoma, Retinal disease, and Diabetic Retinopathy are the five types of Retinopathy. Each category is characterized by specific visual features; for instance, a cataract-affected eye presents with a clouded lens, an eye with glaucoma shows abnormalities in the optic disc, and retinal images associated with retinal disease may exhibit wool patches or variations in optic nerves.

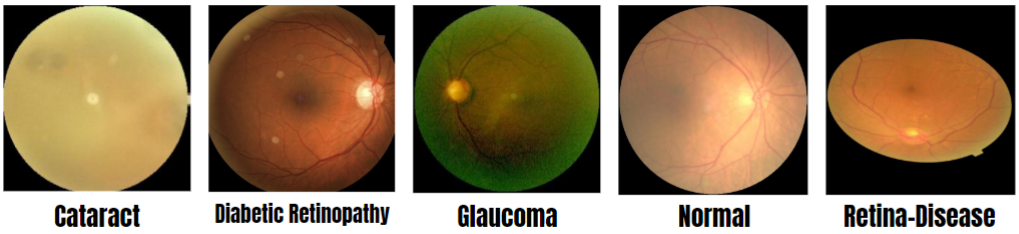


Figure 4.1: Retina images of (a)Cataract, (b) Diabetic retinopathy, (c) Glaucoma (d) Normal (e) Retina Diseases

Using Image annotation tool Roboflow we preprocessed that mentioned data and consisted of 7357 photos that were randomly divided into training ,validation and testing data. around 90% of the images, or 5909 images, were utilized for training, 5%, or 282 images, were used for validation and 5%, or 283 images, were used for testing.Dataset splitting ratio is shown in figure 4.2.1 ,4.2.2 ,4.2.3

4.2 Data Preparation

We use roboflow for our data preparation.Roboflow is an advanced platform utilized for the management and preprocessing of datasets tailored for computer vision applications. We leveraged Roboflow to annotate and prepare our dataset for detecting and classifying five categories of eye diseases. This ensured that our dataset

Disease Category	Test Images
Retina Disease	75
Cataract	68
Normal	55
Glaucoma	44
Diabetic Retinopathy	43
Total	283

Figure 4.2: Data splitting for test.

Disease Category	Validation Images
Retina Disease	74
Cataract	67
Normal	54
Glaucoma	44
Diabetic Retinopathy	43
Total	282

Figure 4.3: Data splitting for validation.

Disease Category	Total Images	Train Images
Retina Disease	1941	1584
Cataract	1753	1400
Normal	1418	1139
Glaucoma	1136	905
Diabetic Retinopathy	1109	881
Total	7357	5909

Figure 4.4: Data splitting for train

was not only well-structured but also augmented to enhance model performance and robustness.

4.3 Annotation process

We used object detection to localize and classify medical images related to eye diseases into five predefined classes:Retina Disease,Cataract,Normal Glaucoma,Diabetic

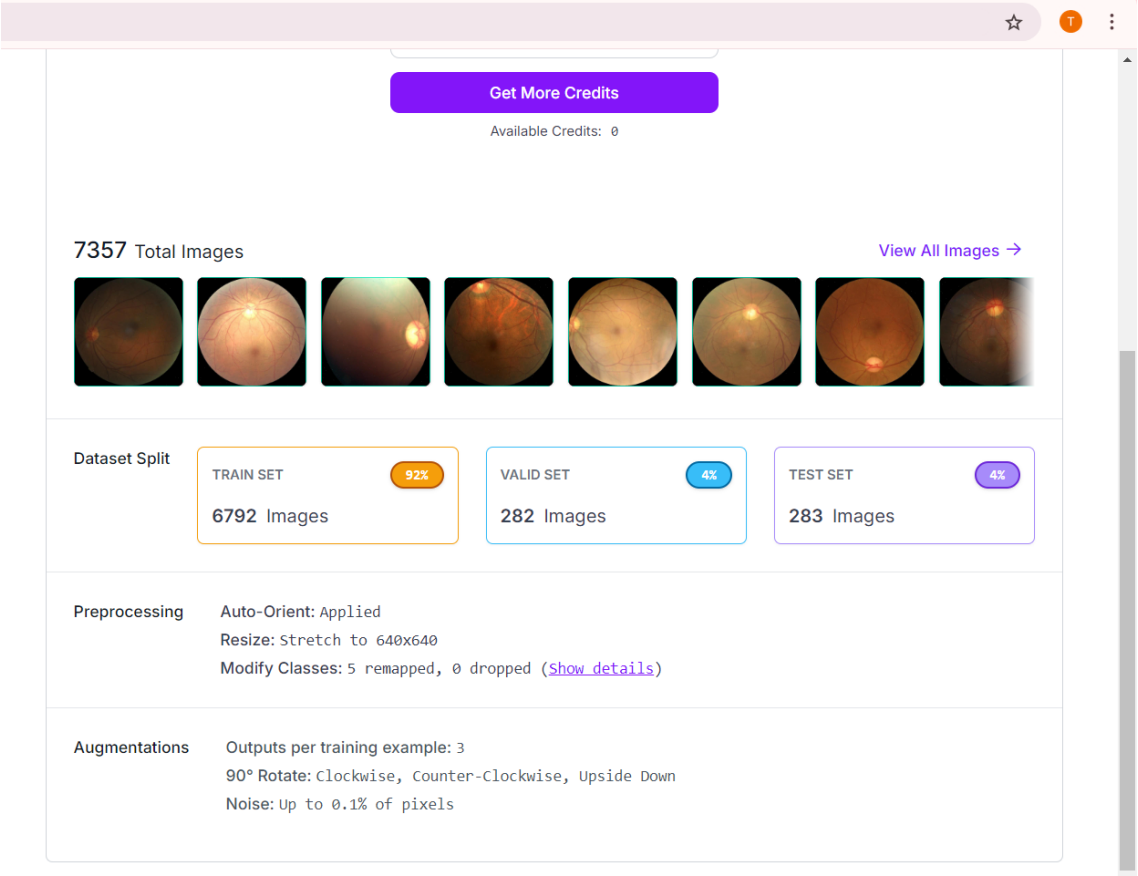


Figure 4.5: overview of annotation.

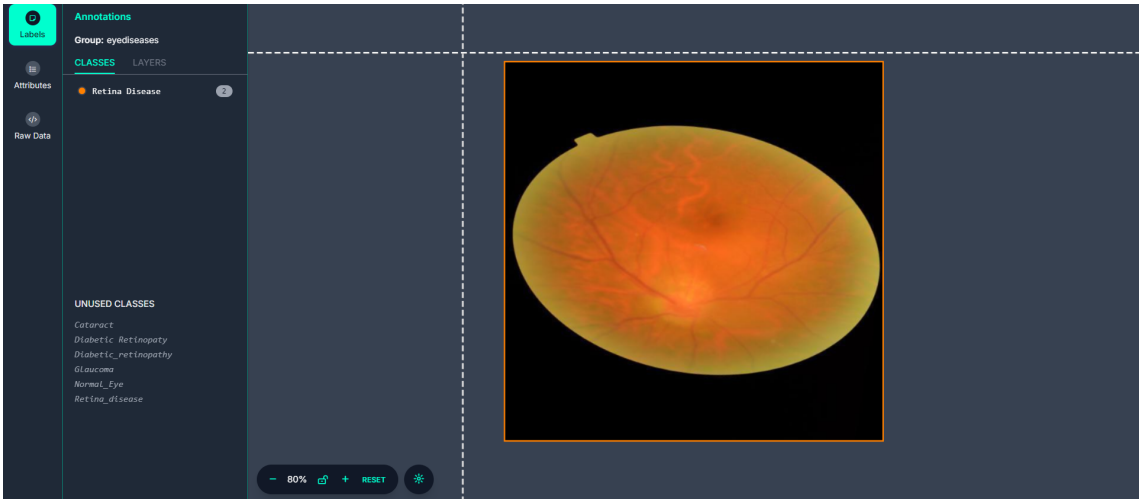


Figure 4.6: Annotated class for the Retina Diseases.

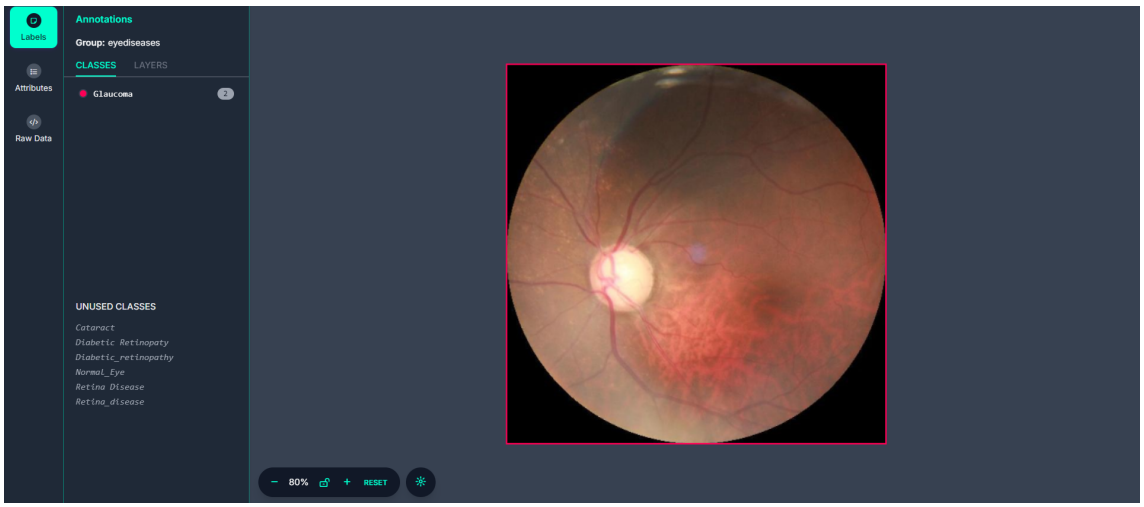


Figure 4.7: Annotated class for the Glaucoma

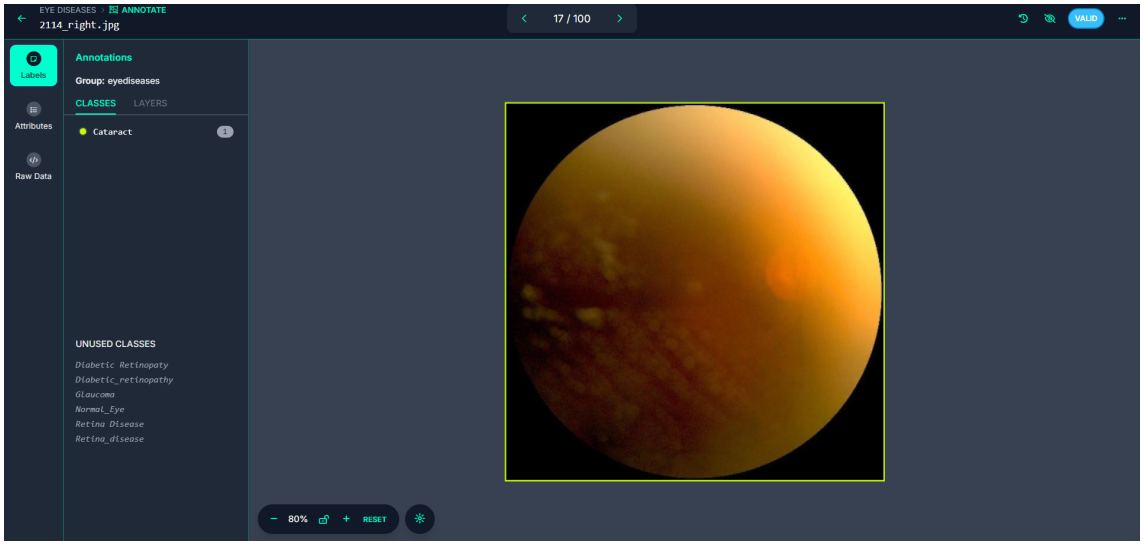


Figure 4.8: annotated class for the cataract

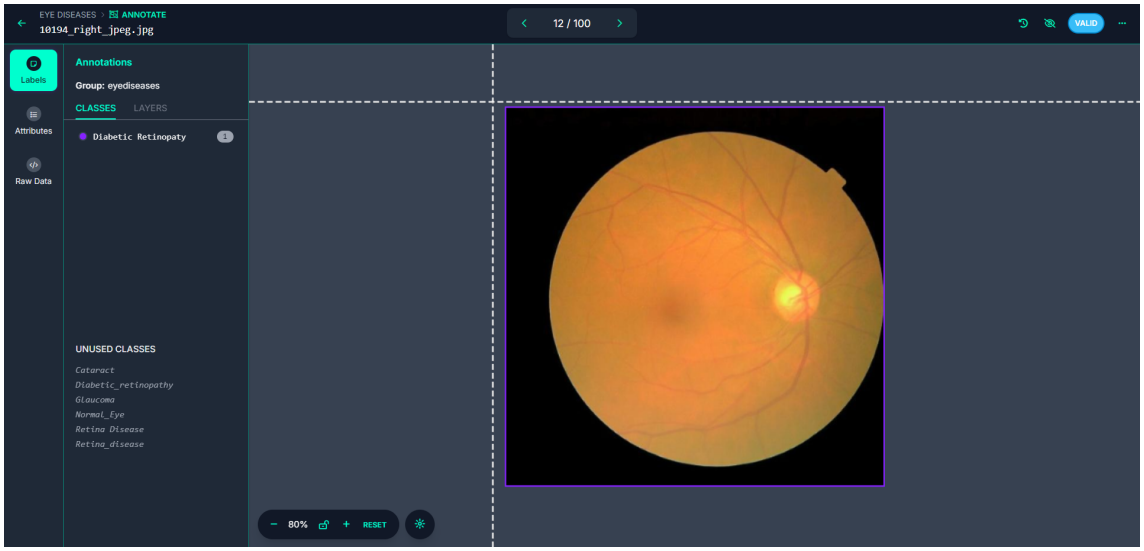


Figure 4.9: annotated class for the Diabetic Retinopathy

4.4 Pre procesing

For the pre processing apply the the Auto-Orient: Applied Resize: Stretch to 640x640 Modify Classes: 5 remapped, 0 dropped

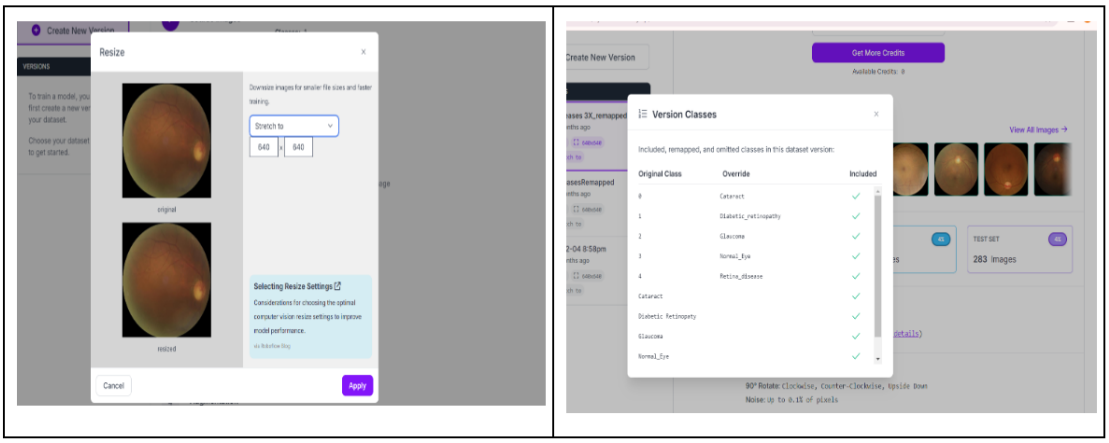


Figure 4.10: pre processing.

4.5 Data Augmentation

Data augmentation is a vital component in preparing our dataset for training machine learning models, especially for object identification and picture classification applications. One of the key motivations for using data augmentation is to expand the useful size of our dataset. In many circumstances, particularly in specialized sectors such as medical imaging, the amount of datasets is restricted. By producing variants of existing photos using techniques like as rotation, flipping, and brightness modifications, we may dramatically increase the dataset's variety, hence improving the model's capacity to learn robust features. Additionally, data augmentation enhances the model's generalization capabilities. By exposing it to a wide range of transformations, the model becomes more resilient to variations it may encounter in real-world scenarios. This is particularly important in medical imaging, where conditions can vary greatly. Moreover, augmentation acts as a form of regularization, helping to prevent overfitting. It ensures that the model does not simply memorize the training examples but instead learns to recognize underlying patterns applicable to different instances of the same class.

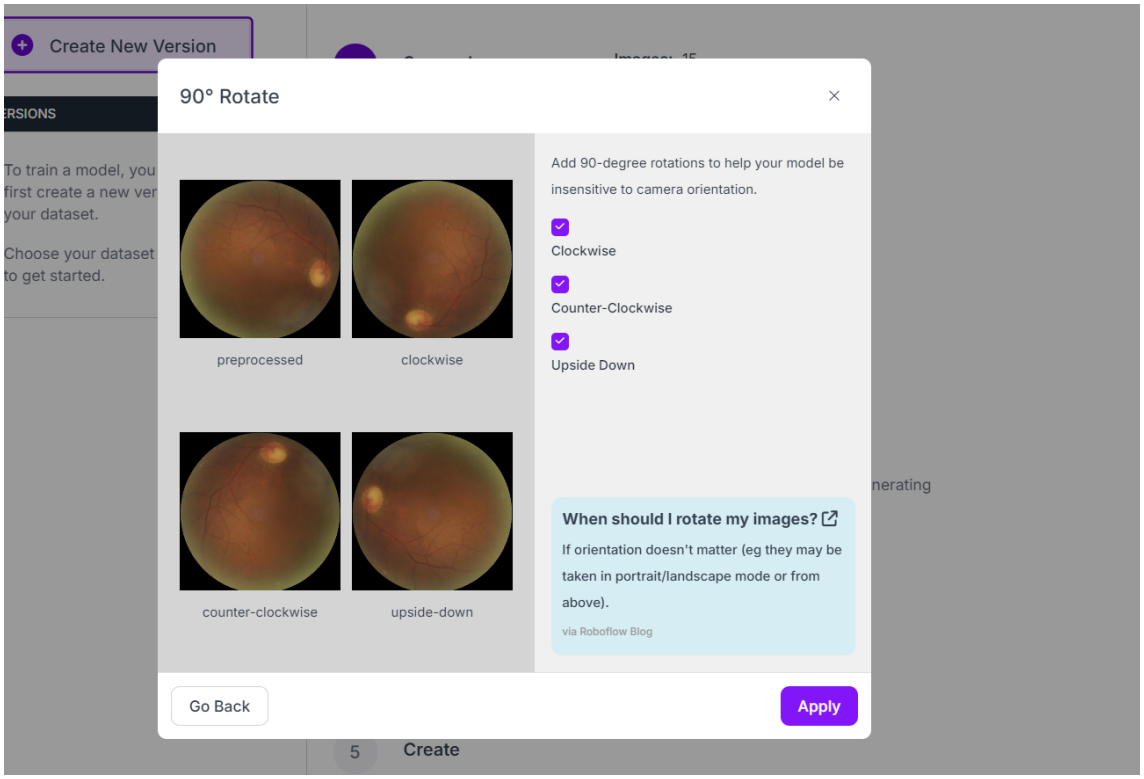


Figure 4.11: augmentation.

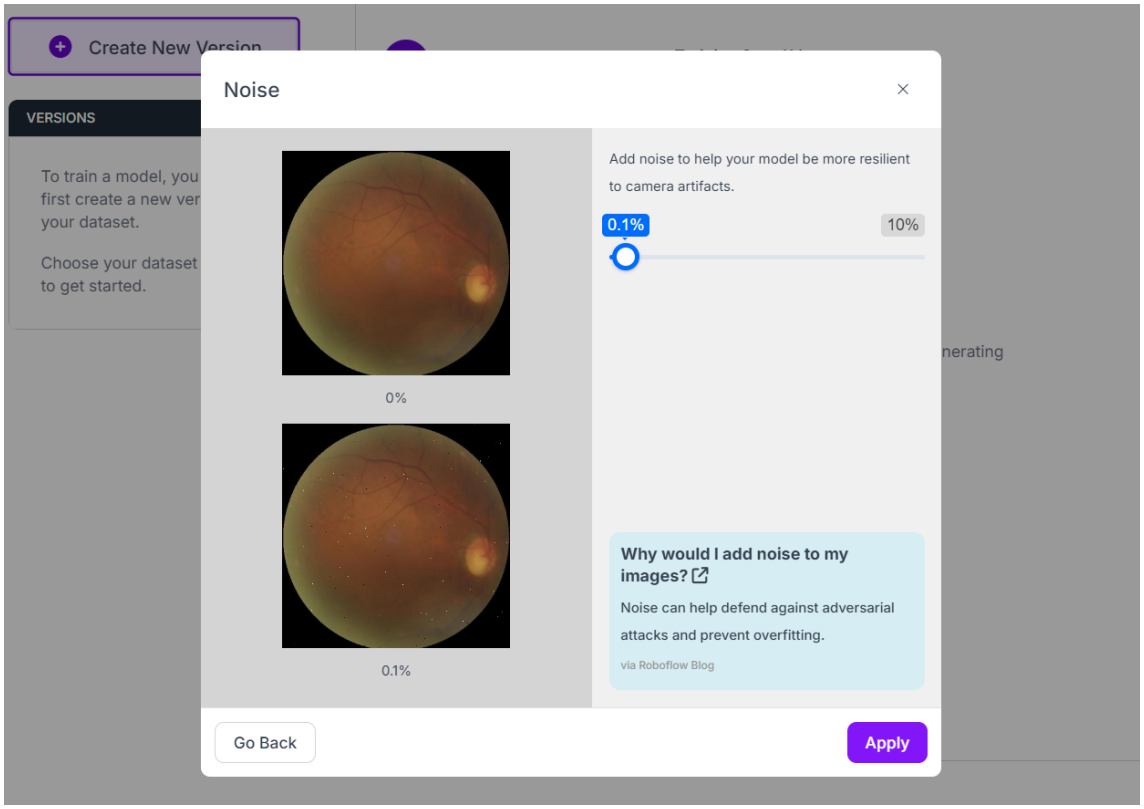


Figure 4.12: pre processing.

Description of the Model

4.6 MODEL(YOLOv5)

In CNN, the feature extraction process producess sets of feature with certain samples and attribute by involving input data scanning with convolutional network.These sets preserve hierarchical information from the input, allowing the proposed model to recognize and interpret complicated patterns and structures critical for EDC(Eye Disease Classification). YOLOv5, a successful object recognition technique, finds and localizes items in fundus images using a single neural network forward pass.
The fundamental architecture of YOLOv5 comprises a deep CNN, with detecting heads extracting features from the input image at different sizes and resolutions. Distinctive feature maps were generated across various layers of the network, capturing nuanced levels of detail and abstraction. YOLOv5 strategically places anchor boxes on the feature map, determining their positions based on predefined dimensions and aspect ratios, as visually represented in Figure 4.3 .

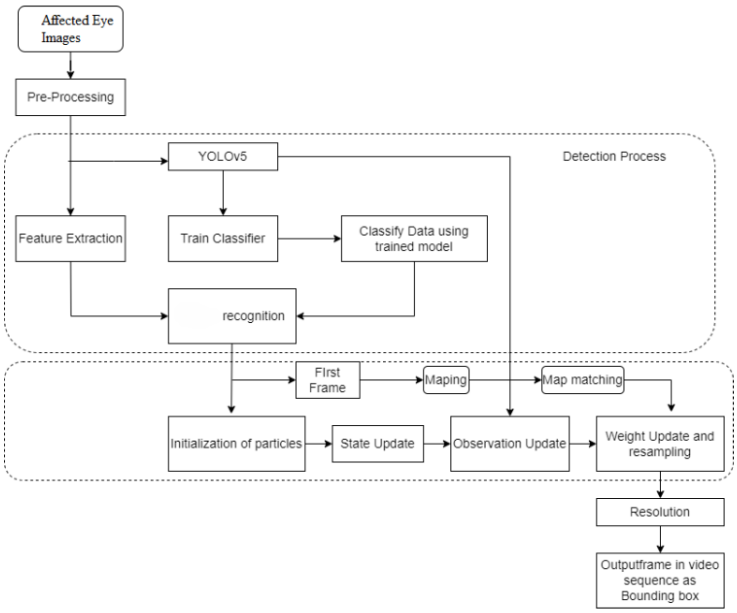


Figure 4.13: Model Structure of Eye Diseases Detection

The YOLO (You Only Look Once) model is renowned for its quick object detection capabilities. Despite its compact size, it delivers high performance and has shown consistent improvement [22-23]. YOLOv5 is the latest version, succeeding YOLOv4, and is considered faster than both YOLOv4 and YOLOv3, according to [21]. YOLOv5 is user-friendly, easy to configure, implement, and test, while still being backed by substantial research. Since its inception in 2015, YOLO has excelled in real-time object identification for both video and image objects.The envisioned utilization of YOLOv5 is focused on extracting features from input images to predict diseases. Subsequently, a prediction system employs these features to delineate bounding boxes around objects and identify their specific classes.

YOLOv5 is the first member of the YOLO family to be written in PyTorch framework and it is comparatively lightweight and easy to use. It comes in many versions - nano(n), small(s), medium(m), large(l), and extra large(x). The relationship between speed and accuracy is inverse in these versions. So, the YOLOv5n is the fastest and the YOLOv5x provides the highest accuracy among these versions.

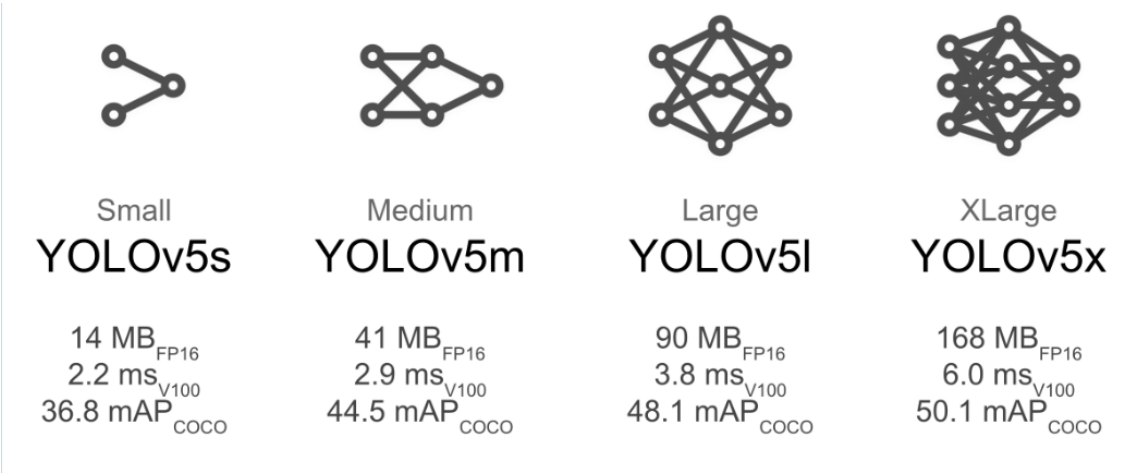


Figure 4.14: Speed vs Accuracy in different YOLOv5 versions.

YOLOv5 is a predictive tool for eye diseases. This process involves taking an input image and processing its characteristics through a prediction system. Then, we outline bounding boxes around the image and predict the corresponding classes [23-24].
This object detector integrates class labels and corresponding bounding boxes from different networks. In Figure 4.6, three essential components for the YOLO network are depicted. CSPDarknet serves as the backbone. PANet acts as the neck, and the YOLO Layer functions as the head. The data undergoes feature extraction in CSPDarknet. Then, it is sent to PANet for feature fusion. As illustrated in Figure 4.5 , the YOLO output per unit includes class, score, position, and size.

4.7 The Whale Optimization Algorithm (WOA)

YOLOv5 imitates the cooperative hunting action of humpback whales by combining aspects of exploration and exploitation. Whales explore by moving randomly over a predefined search space, which encourages the discovery of previously undiscovered regions within the domain. Whales in the exploitation phase modify their positions based on where other whales are in relation to the optimal solution. The goal of this stage is to direct the algorithm toward convergence to the best answer. Whales relocate based on a variety of factors, including their fitness levels, prior placements, and the positions of other whales. When the algorithm reaches the maximum number of iterations or meets predefined stopping criteria, the algorithm terminates. The best way to solve the optimization problem depends on the whale's response.

The algorithm adeptly combines exploration and exploitation techniques to navigate intricate optimal landscapes and promptly pinpoint the most effective feature set. The success of this approach is contingent upon the configuration of parameters and the intrinsic characteristics of the optimization problem. Within the search space, a group of potential solutions (whales) is introduced, and their positions evolve based on their closeness to the optimal solution. During the exploration phase, the authors employ the Levy Flight and Wavelet mutation strategy, enhancing the Whale Optimization Algorithm's (WOA) performance and effectively contributing to the algorithm's proficiency in finding optimal solutions. Basically wavelet mutation strategy uses controlled randomness to enhance the exploration capability of the Whale Optimization Algorithm. The goal is to increase the algorithm's ability to escape local optima by generating slightly modified solutions. Also, Levy Flight is a stochastic search strategy , inspired by animal foraging patterns, where movements alternate between short exploratory steps and occasional long jumps. In the algorithm, long jumps increase the probability of escaping local optima and short jumps ensure local search refinement after large steps. Together, these techniques significantly improve WOA's robustness and convergence speed.

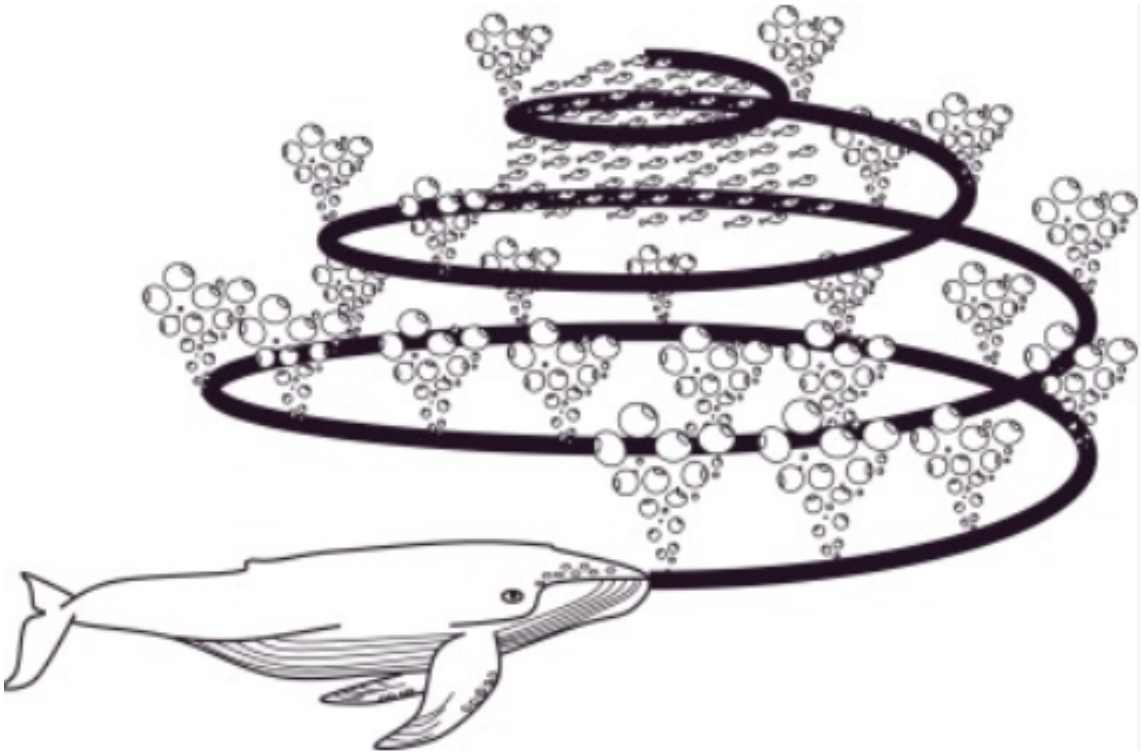


Figure 4.15: The Whale Optimization Algorithm (WOA).

4.8 Architecture of YOLOv5

The YOLO family of models is structured around three key building blocks: the Backbone, Neck, and Head.

YOLOv5 Backbone

YOLOv5 employs CSPDarknet as the backbone for feature extraction from images [1] and this CSPDatknet is a modified version of Darknet Architecture originally used in YOLOv3. It is basically based on DenseNet which was designed to reduce vanishing gradient problem and strengthen feature propagation at the same time which reuses features and reduces network parameters [30]. The inclusion of CSP(Cross Stage Partial) blocks improves the backbone's ability to balance accuracy and computational costs. Figure 4.6 illustrates the architecture of the backbone of YOLOv5.

YOLOv5 Neck

YOLOv5 uses components like feature pyramid FPN, and path aggregation structure PANet. FPN top-down network uses up-sampling to transfer and fuse information to predict the feature map and PANet bottom-up transfer of localization of information and fusion of information from different network layers in the backbone helps to enhance the detection capabilities [1]. The neck basically takes feature maps from the backbone and aggregates and refines them which improves multilayer detection.

YOLOv5 Head

This portion of the architecture uses anchor boxes to identify objects effectively. It is outfitted with layers meant to produce predictions. It uses the feature maps generated by the head and processed by the neck to output the bounding box, class possibilities and confidence scores for the detected object.

In addition, YOLOv5 makes specific choices during training [29]

Activation and optimization: It uses leaky ReLU and sigmoid activation functions, allowing flexibility with optimizer options such as SGD and ADAM.

YLoss function: For efficient model training, YOLOv5 uses binary cross- entropy with logits loss

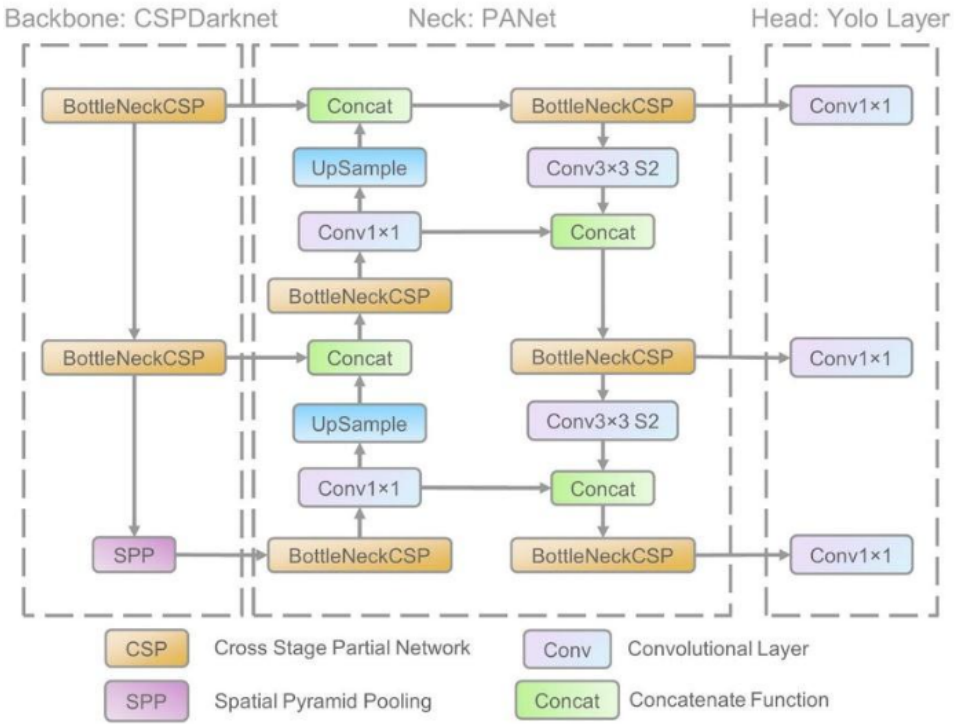


Figure 4.16: Backbone architecture of Yolov 5.

4.9 Implementation

The suggested model was trained using the Eye Diseases Classification dataset (EDC). The Eye Diseases Classification dataset (EDC) was made using primary and secondary [18] datasets. The primary dataset was collected from the National Institute of Ophthalmology, Dhaka, Bangladesh. And the secondary dataset was collected from multiple repositories. The combined dataset was classified into five categories: Normal, Cataract, Glaucoma, Retinal disease, and diabetic retinopathy. Based on the picture classifications in the EDC dataset, the proposed model divides images into these five groups. The EDC datasets provide ground truth information. The important attributes of these datasets are summarized in Figure 4 , and sample photos from the EDC datasets are shown in Figure 4.1.
The YOLOv5 model developed for this article is structured into two parts: a training model and a detection model, as illustrated in Fig. 3. The training model utilizes 6792 images. On the other hand, the detection model analyzes input images from the Eye Diseases Detection dataset and generates predictions for five classes: Normal, Cataract, Glaucoma, Retinal disease, and Diabetic retinopathy.
We conducted multiple tests to assess the performance of the proposed model. The tests were conducted under various experimental configurations. Model instructions were modified by altering several network parameters. The complete dataset was divided into 5The input images were first resized to the predefined input size of the Backbone network for feature extraction. Then this feature set was aggregated and refined in the neck and then the head uses an anchor box to identify objects and outputs the bounding box, class probabilities and confidence score. After the forward pass, YOLOv5 calculates the loss functions. Once they are calculated, YOLOv5 uses back propagation techniques to adjust the weights, minimizing the loss. The model's size was determined for the identification of the object using the parameters in the weight file. The mAP@.5 was used to calculate the average AP across all categories. The mAP@.5:.95 was also calculated. The performance of the model was evaluated using Precision, Recall and Accuracy. The following equations show their formulas-
Precision=(True Positives (TP)/(True Positives (TP)+False Positives (FP)))*100

Recall=(True Positives (TP)/(True Positives (TP)+False Negatives (FN)))*100

Accuracy=(True Positives (TP)+True Negatives (TN))/(True Positives (TP)+True Negatives (TN)+False Positives (FP)+False Negatives (FN)) * 100
The loss function is given below

$$\text{loss } = l_{bax} + i_{cls} + l_{obj}$$

The bounding box regression is defined in the following way-

$$l_{box} = \lambda_{\text{coord}} \sum_{i=0}^{\infty} \sum_{j=0}^{D} I_{i,j}^{ojb} bj \left(2 - w_i \times h_i\right)$$
$$\left[\left(x_i - x\wedge_i^j\right)^2 + \left(y_i - y\wedge_i^j\right)^2 + \left(w_i - w\wedge_i^j\right)^2 + \left(h_i - h\wedge_i^j\right)^2\right]$$

The classification loss function is denied in the following way

$$\mathrm{l}_{cls} = \lambda_{\text{class}} \sum_{i=0j}^{s^2} \sum_{=0}^{B} I_{i,j}^{obj} \sum_{c \in c1asses} p_i(c) \log\left(\hat{p}_l(c)\right).$$

And the confidence loss definition is given below-

$$l_{obj} = \lambda_{noobj} \sum_{i=0}^{s^*} \sum_{=0}^{B} I_{i,j}^{noobj} \left(c_i - c\wedge_l\right)^2$$
$$+$$
$$\lambda_{obj} \sum_{i=0}^{s^2} \sum_{j=0}^{B} I_{i,j}^{obj} \left(c_i - c\wedge_l\right)$$

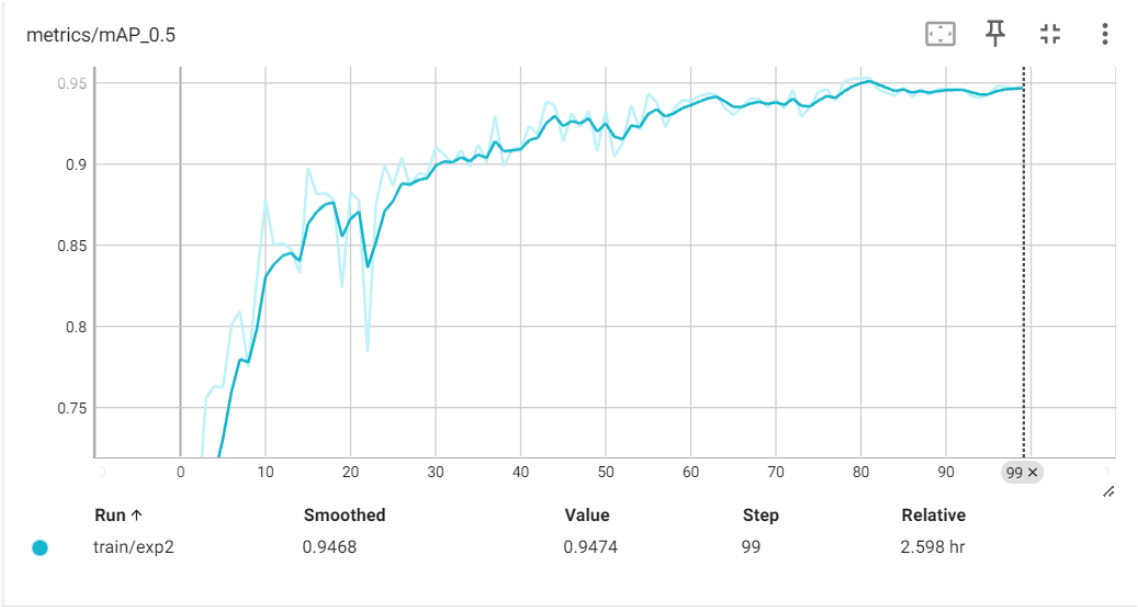


Figure 4.17: mAP@.5 of the model.

The model used the following loss functions:
We also carried out additional experiments. We altered the pre-trained model, such as YOLOv7 and Inception- ResNet50.

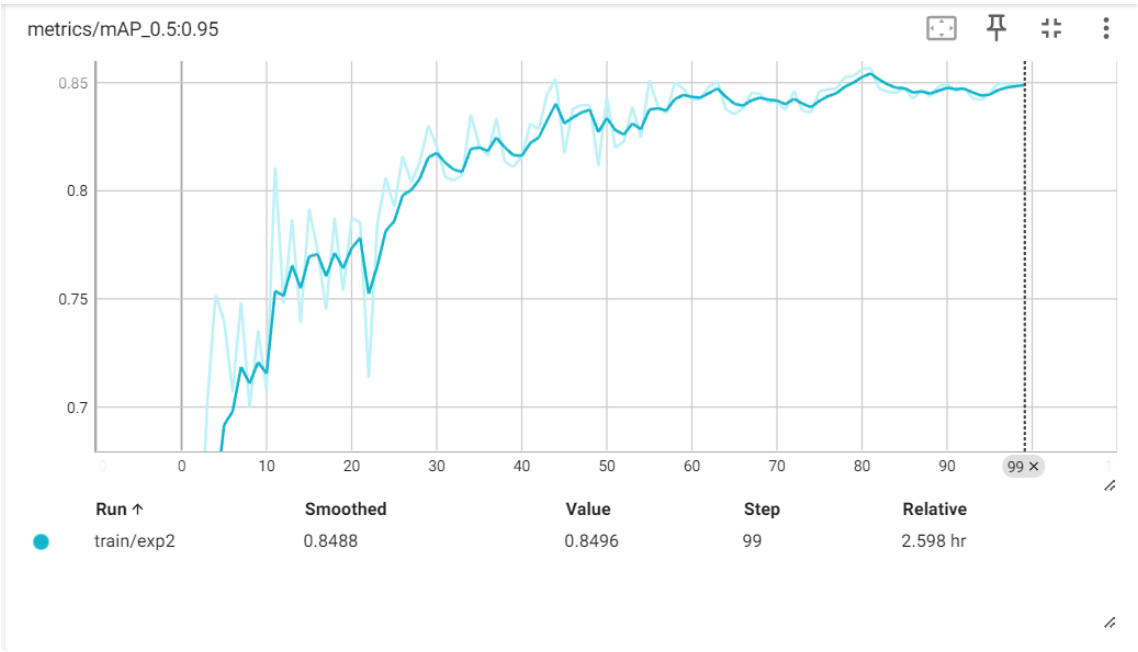


Figure 4.18: mAP@.5:.95 of the model.

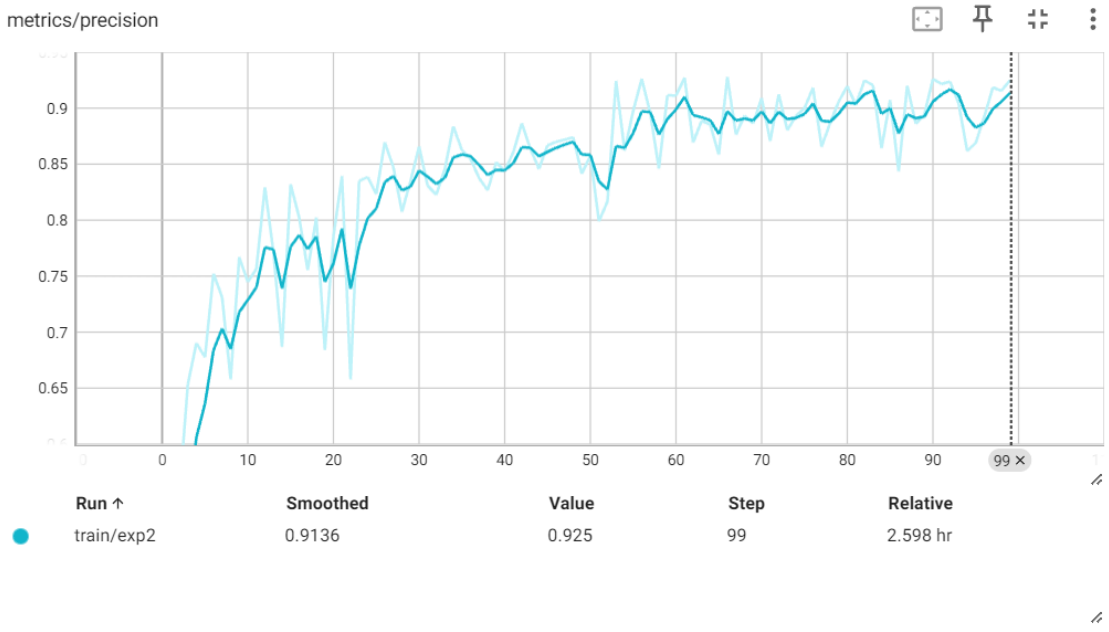


Figure 4.19: precision of the map .

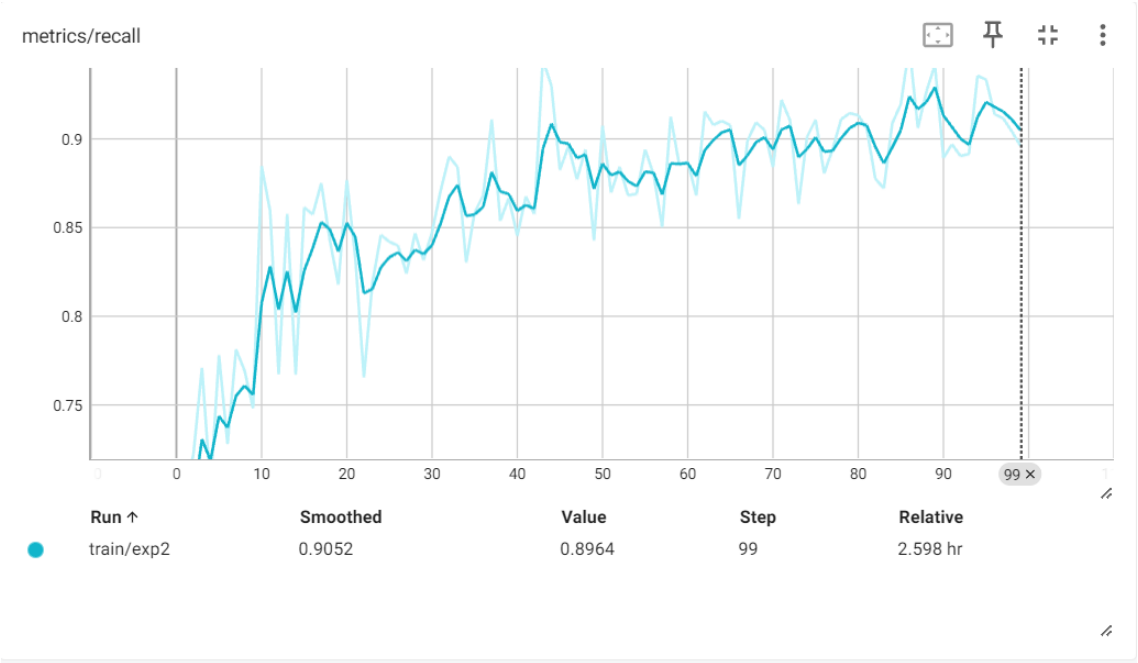


Figure 4.20: Recall of the map.

4.10 Preliminary Analysis

Our model was trained using the training dataset, including original and augmented photos. A validation test was conducted to evaluate the YOLOv5 and YOLOv7 models' generalization capabilities. The network demonstrated satisfactory convergence during the training stages, as shown graphically in Figure 4.24.
The model was developed on a Windows 10 system with 16 GB RAM, an Intel Core i7 processor, and an NVIDIA GeForce RTX 3050 graphics card, using Python 3.8.3. The dataset was divided into training (90%), validation (5%), and testing (5%) sets. The model performed impressively throughout training, achieving a training loss of 0.2%. and a high accuracy of 98.74%. During validation, the model maintained a strong performance, with a loss of 2.254%. and an accuracy of 92.1%.. Additionally, the model displayed a sensitivity of 87.1 % and a perfect specificity of 100%., reflecting its proficiency in correctly identifying both positive and negative instances. This balanced performance suggests high accuracy and reliability.

		Retina Disease	Cataract	Normal	Glaucoma	Diabetic retinopathy
Yolov5	Precision	0.883	0.699	0.859	0.732	0.969
	Recall	0.973	0.875	0.932	0.561	0.955
	mAP	0.949	0.841	0.911	0.711	0.941
Yolov7	Precision	0.29	0.637	0.385	0.285	0.649
	Recall	0.973	0.845	0.658	0.342	0.756
	mAP	0.583	0.828	0.714	0.287	0.714

Figure 4.21: Training accuracy about all classes of the dataset.

4.11 Result of YOLOv5

After evaluation, the test dataset figure shows that the average detection accuracy of all five classes by using the YOLOv5 model is 95%. at mAP50 and 85%. at mAP50-95. Where Cataract got the highest accuracy rate with 97.8%., Diabetic Retinopathy got 95.4%., Glaucoma got 88.7%., Normal eye got 97.1%.and Retinal Disease got 97.6%.

```
Model summary: 212 layers, 20869098 parameters, 0 gradients, 47.9 GFLOPs
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100%|██████████| 4/4 [00:04<00:00,  1.16s/it]
                   all        282        285      0.903      0.905      0.953      0.857
              Cataract        282         61      0.887      0.984      0.978       0.88
   Diabetic Retinopaty        282         43       0.93      0.932      0.954      0.857
              Glaucoma        282         38      0.791      0.799      0.887      0.795
            Normal_Eye        282         61      0.932      0.905      0.971      0.874
         Retina Disease        282         82      0.974      0.908      0.976      0.878
Results saved to runs/train/exp2
```

Figure 4.22: Result of yolov5 training

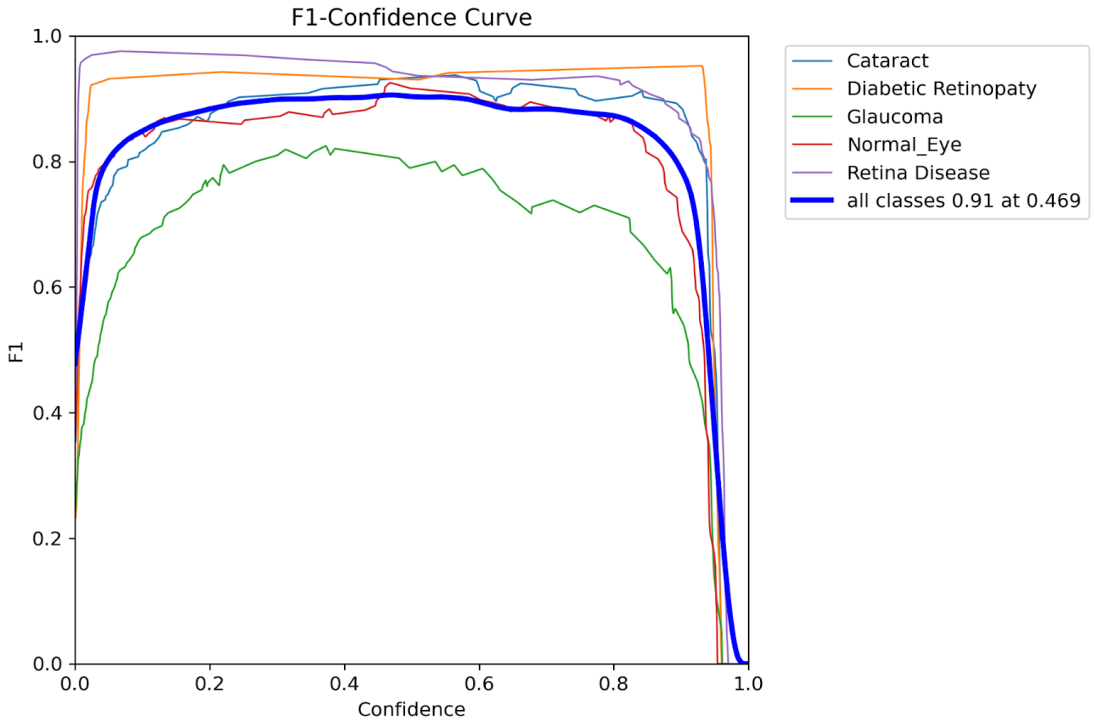


Figure 4.23: *F1_curve*

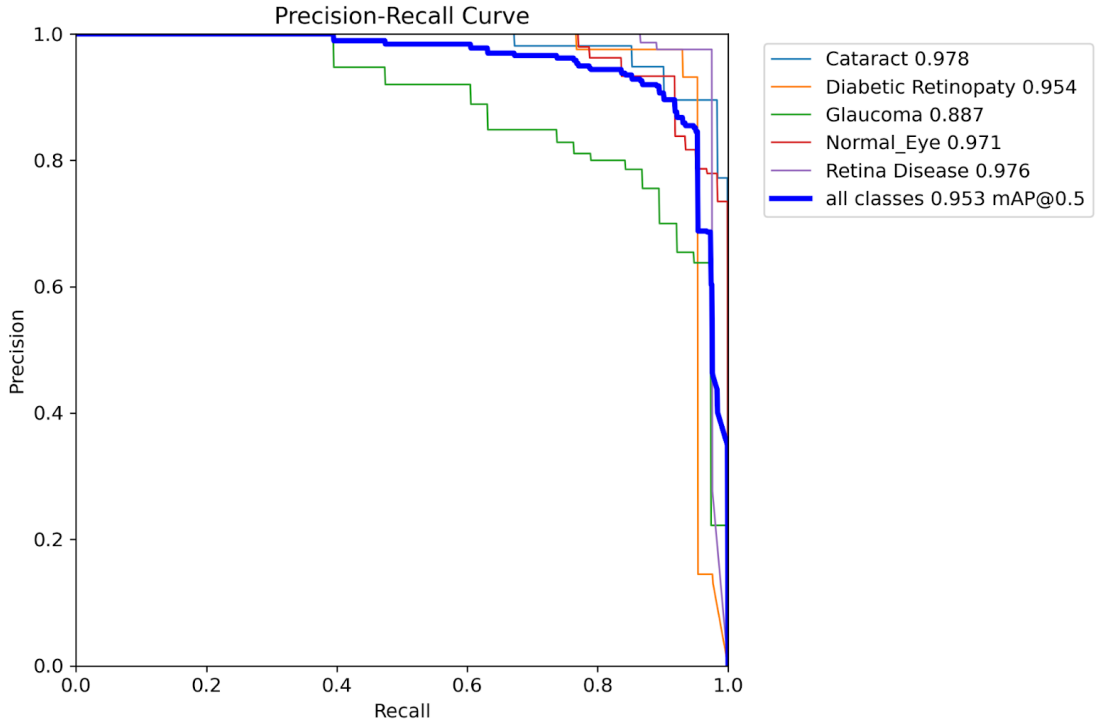


Figure 4.24: *PR_curve*

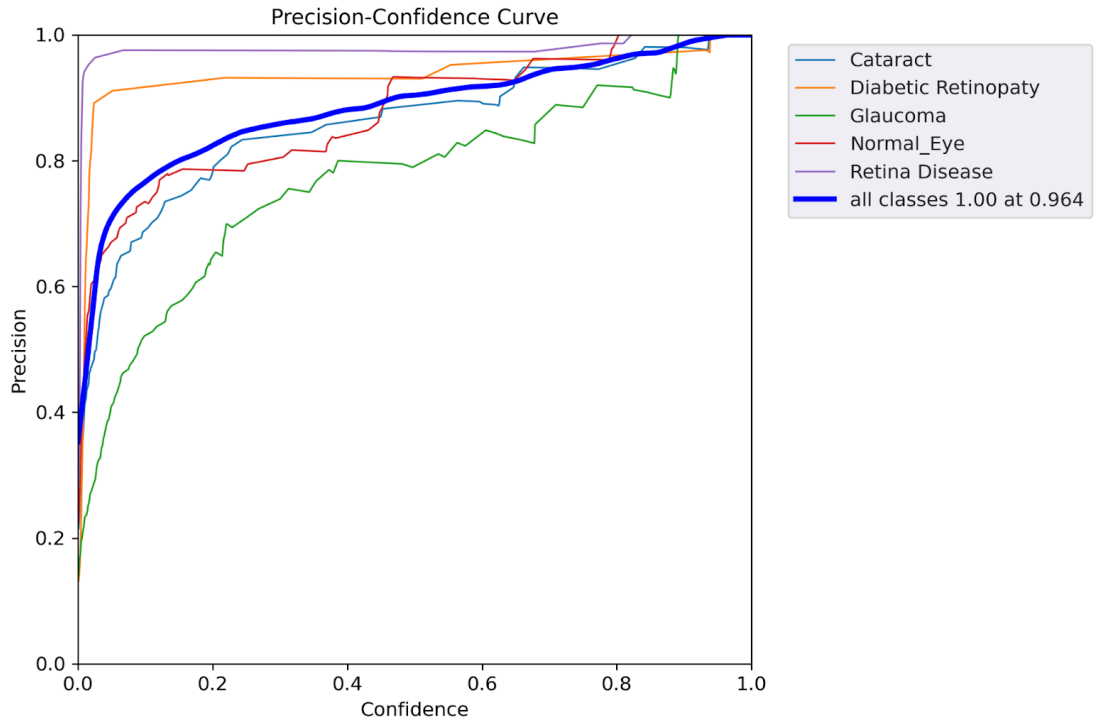


Figure 4.25: *P_curve*

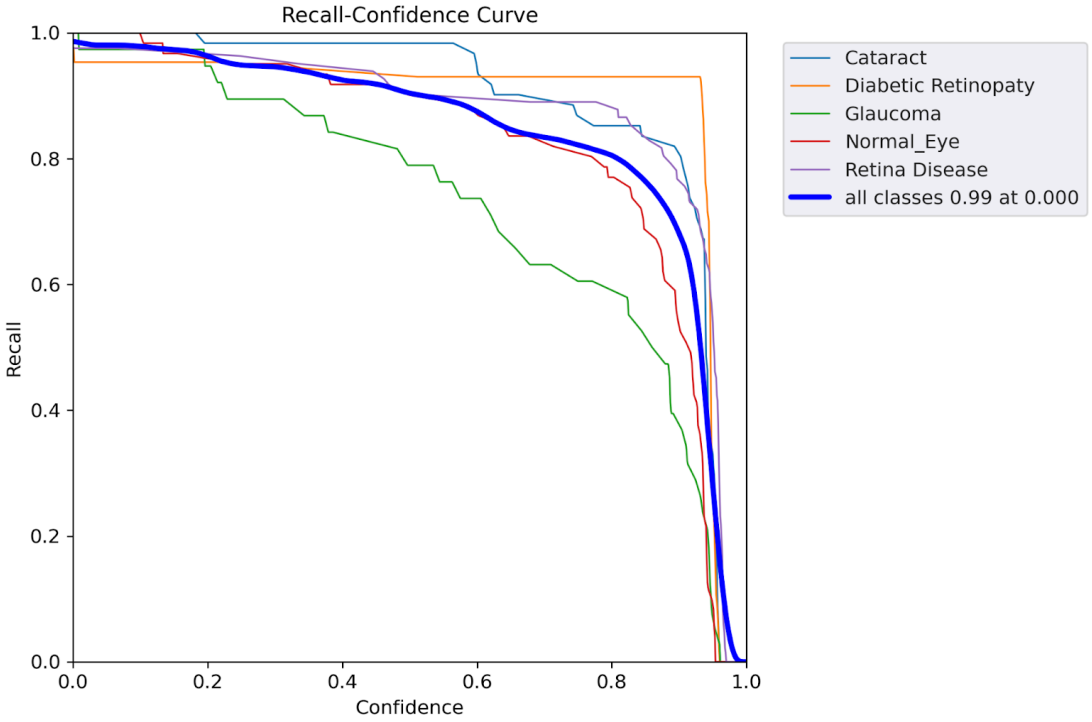


Figure 4.26: *R_curve*.

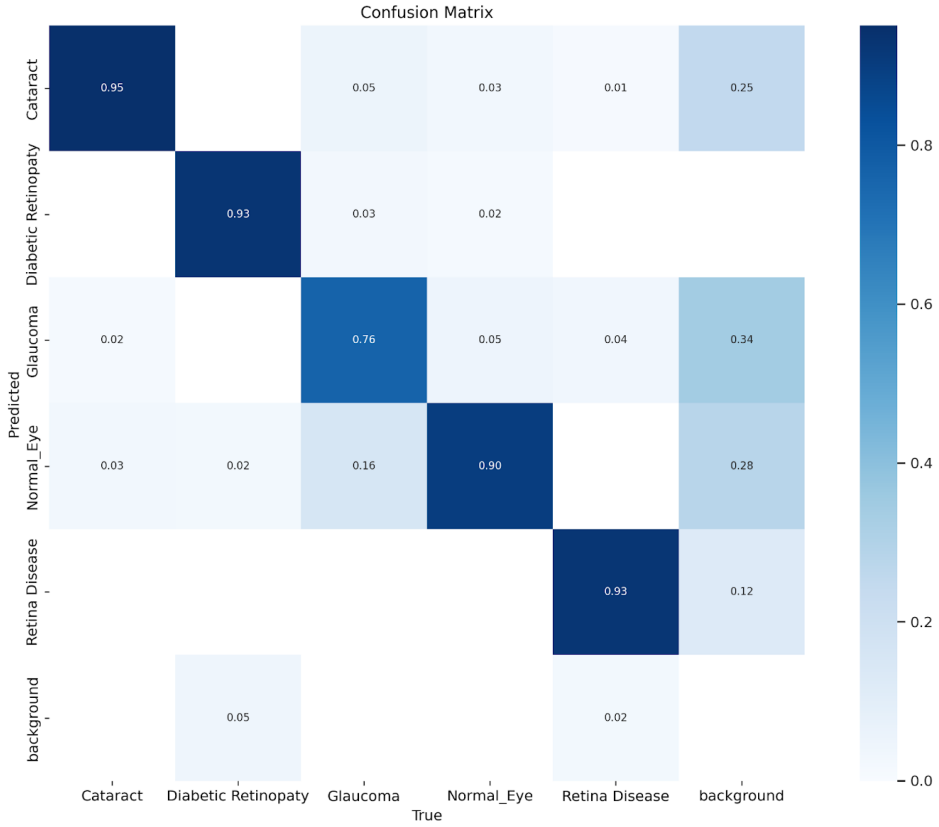


Figure 4.27: Confusion Matrix.

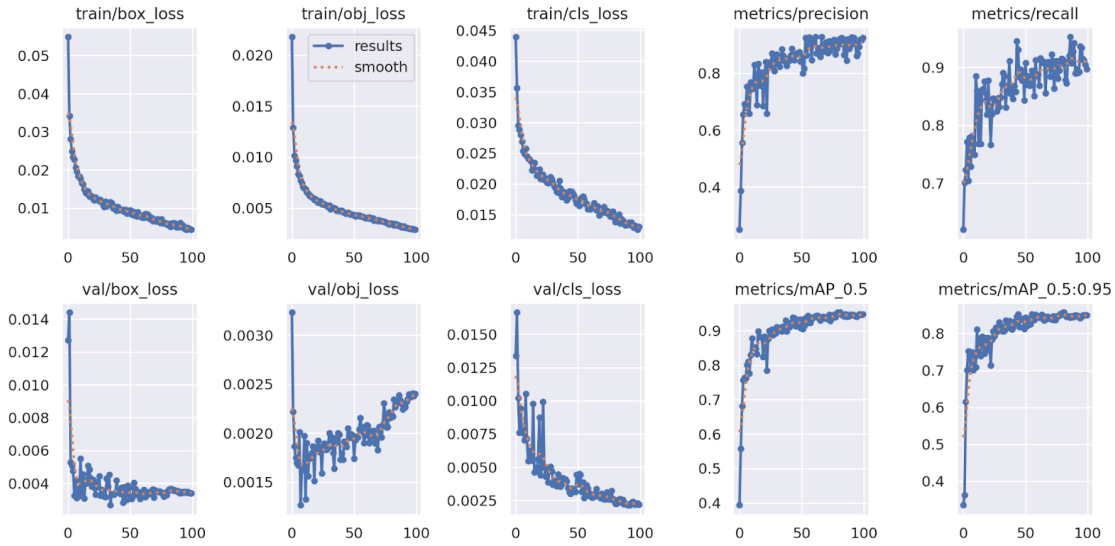


Figure 4.28: Result.

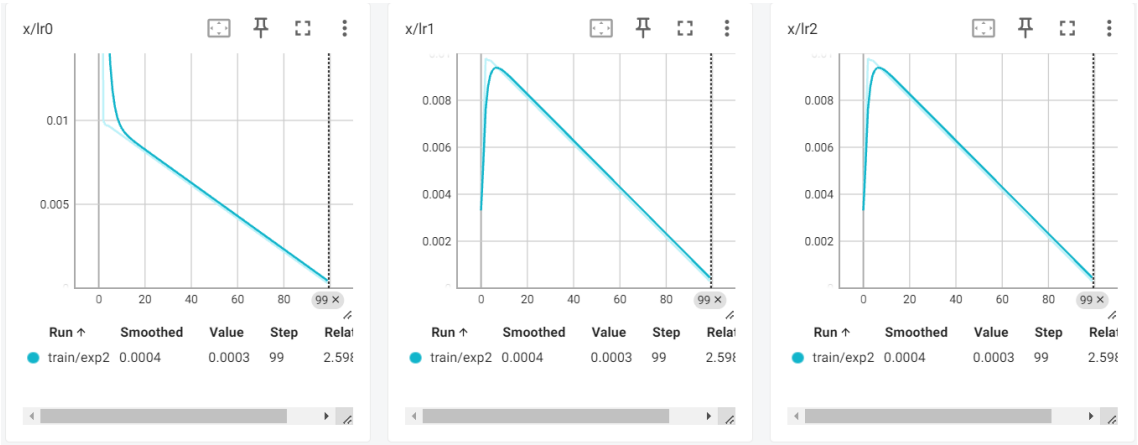


Figure 4.29: $x/Ir0, x/Ir1, x/Ir2$.

4.12 Result of YOLOv7

After evaluation, the test dataset figure shows that the average detection accuracy of all five classes by using the YOLOv7 model is 60% at mAP@.5 and 56 at mAP@.50:.95. Where Cataract got the highest accuracy rate with 86%, Diabetic Retinopathy got 73.2%, Glaucoma got 29.7%, Normal eye got 39.8% and Retinal Disease got 71.1%

Class	Images	Labels	P	R	mAP@.5	mAP@.5:.95: 100% 9/9 [00:08<00:00, 1.07it/s]
all	279	282	0.449	0.715	0.6	0.56
Cataract	279	71	0.637	0.845	0.861	0.828
Diabetic_retinopathy	279	41	0.649	0.756	0.732	0.714
Glaucoma	279	38	0.285	0.342	0.297	0.287
Normal	279	59	0.385	0.658	0.398	0.389
Retina_Disease	279	73	0.29	0.973	0.711	0.583

100 epochs completed in 3.709 hours.

Figure 4.30: Result of yolov7 test dataset

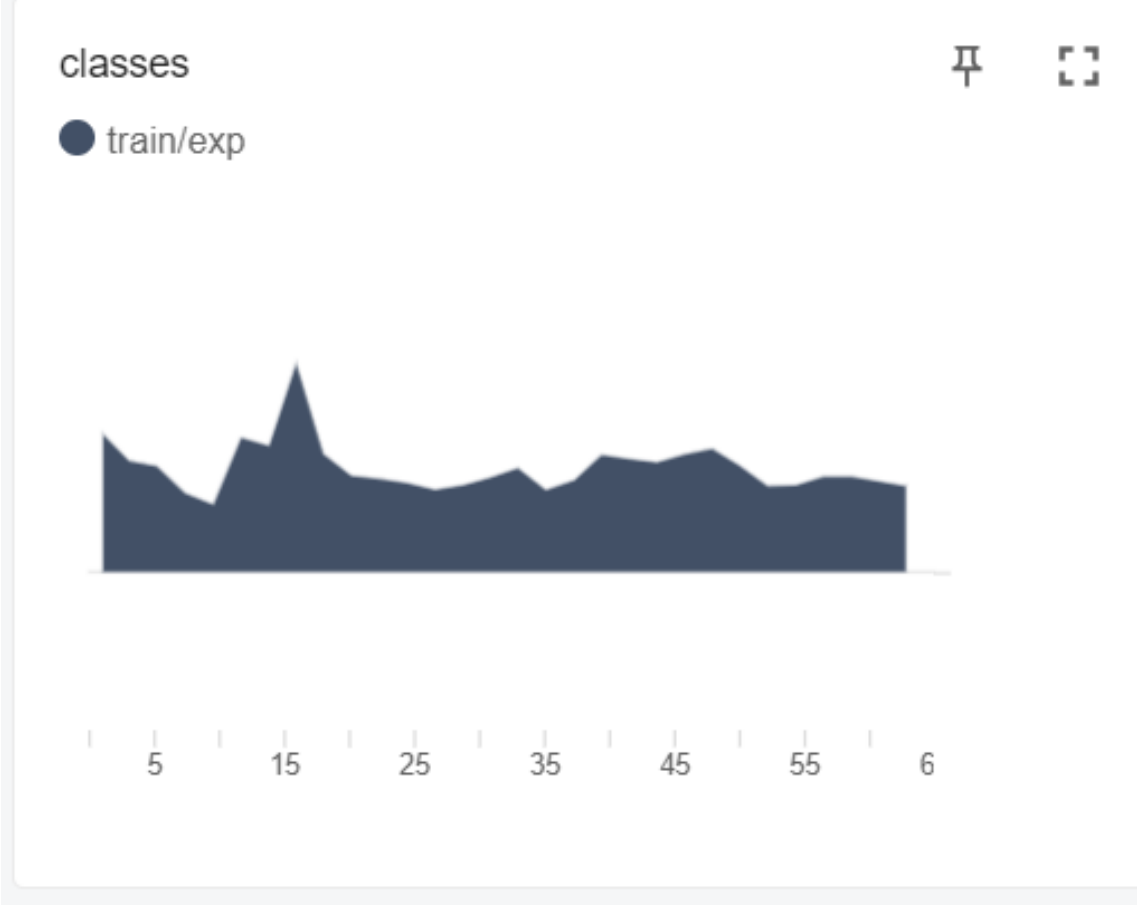


Figure 4.31: The class(train/exp) curve.

4.13 Final Result

The effectiveness of our proposed model was benchmarked against other traditional models on the same dataset: **YOLOv5:** Achieved a mAP of 0.951 and precision of 0.926, outperforming other models in accuracy and precision. **YOLOv7**: Reached a precision of 0.449 and mAP of 0.544. **-Inception-ResNet50**: Recorded a validation accuracy of 90%., precision of 0.63, and mAP of 0.901.
These comparisons highlight YOLOv5's superior performance, making it the most suitable model for this task. The suggested model, with optimized parameters via AO, effectively reduced computational requirements while maintaining high accuracy, underscoring its capability in classifying eye disease images.

Conclusion and Future Work

This article describes the development of an automated diagnostic approach for Normal, Cataract, Glaucoma, Retinal disease, and Diabetic retinopathy using pre-trained models, such as InceptionResNet50, YOLOv5, and YOLOv7. In particular, the Yolov5 model achieves impressive performance metrics with a low training error of 0.2% and a high training accuracy of 98.74%, demonstrating substantial skills in recognizing eye illnesses. During validation, the model continues to perform well, displaying a validation loss of 2.254% and an accuracy of 92.1%.In resource-constrained nations with a scarcity of trained ophthalmologists, the article highlights the potential usefulness of an automated cataract diagnostic system.Such solutions could improve healthcare accessibility, cut screening time and costs for both patients and ophthalmologists and aid in early diagnosis. Future work could concentrate on enhancing model accuracy by using larger and more complicated datasets, experimenting with other image-processing methods for more accurate learning, and possibly establishing a user-friendly website for global accessibility.

Reference

- 1. Rachburee, N.,& Punlumjeak, W. (2021). An assistive model of obstacle detection based on deep learning: YOLOv3 for visually impaired people. *International Journal of Electrical & Computer Engineering*, 11(4).
- 2.Heravi, E. J., Aghdam, H. H., & Puig, D. (2016, September). Classification of foods using spatial pyramid convolutional neural network. In *CCIA* (pp. 163–168).
- 3. Amara, J., Bouaziz, B., & Algergawy, A. (2017, March). A deep learning-based approach for banana leaf diseases classification. In *BTW (workshops)* (Vol. 266, pp. 79–88).
- 4. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419.
- 5.Goatman, K. A., Cree, M. J., Olson, J. A., Forrester, J. V., &Sharp, P. F. (2003). Automated measurement of microaneurysm turnover. *Investigative Ophthalmology & Visual Science*, 44(12), 5335–5341.
- 6.Cree, M. J., Gamble, E., & Cornforth, D. J. (2005). Colour normalisation to reduce inter-patient and intra-patient variability in microaneurysm detection in colour retinal images. *Proceedings of Medical Image Understanding and Analysis Conference*.
- 7. Li, H., & Chutatape, O. (2004). Automated feature extraction in color retinal images by a model-based approach. *IEEE Transactions on Biomedical Engineering*, 51(2), 246–254.
- 8.Acharya, U. R., Kannathal, N., Ng, E. Y. K., Min, L. C., & Suri, J. S. (2006, August). Computer-based classification of eye diseases. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6121–6124). IEEE.
- 9. Raja, J., Shanmugam, P., & Pitchai, R. (2021). An automated early detection of glaucoma using support vector machine based visual geometry group 19 (VGG-19) convolutional neural network. *Wireless Personal Communications*, 118, 523–534.
- 10. Patil, N., Patil, P. N., & Rao, P. V. (2021). Convolution neural network and deep-belief network (DBN) based automatic detection and diagnosis of glaucoma. *Multimedia Tools and Applications*, 80(19), 29481–29495.
- 11.Abdel-Hamid, L. (2020). Glaucoma detection from retinal images using statistical and textural wavelet features. *Journal of Digital Imaging*, 33(1), 151–158.

- 12. Yang, M., Yang, J. J., Zhang, Q., Niu, Y., & Li, J. (2013, October). Classification of retinal image for automatic cataract detection. In *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)* (pp. 674–679). IEEE.

- 13.Haleem, M. S., Han, L., Van Hemert, J., & Li, B. (2013). Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: A review. *Computerized Medical Imaging and Graphics*, 37(7–8), 581–596.

- 14. Barros, D. M., Moura, J. C., Freire, C. R., Taleb, A. C., Valentim, R. A., & Morais, P. S. (2020). Machine learning applied to retinal image processing for glaucoma detection: Review and perspective. *Biomedical Engineering Online*, 19, 1–30.

- 15.Rajalakshmi, R., Subashini, R., Anjana, R. M., & Mohan, V. (2018). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*, 32(6), 1138–1144.

- 16.Xiong, L., Li, H., & Zheng, Y. (2014, June). Automatic detection of glaucoma in retinal images. In *2014 9th IEEE Conference on Industrial Electronics and Applications* (pp. 1016–1019). IEEE.

- 17. Gardner, G. G., Keating, D., Williamson, T. H., & Elliott, A. T. (1996). Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool. *The British Journal of Ophthalmology*, 80(11), 940.

- 18.Retina Dataset, Kaggle. Available at: https://www.kaggle.com/datasets/jr2ngb/cataractda

- 19.Jocher, G., Stoken, A., Chaurasia, A., Borovec, J., Kwon, Y., Michael, K., Liu, C., Fang, J., Abhiram, V., & Skalski, S. P. (2021). Ultralytics/yolov5: V6.0—YOLOv5n 'Nano' models, Roboflow integration, TensorFlow export, OpenCV DNN support. *Zenodo*.

- 20.Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10781–10790).

- 21. Su, J., Liu, C., Coombes, M., Hu, X., Wang, C., Xu, X., Li, Q., Guo, L., & Chen, W. H. (2018). Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Computers and Electronics in Agriculture*, 155, 157–166.

- 22. Thuan, D. (2021). Evolution of YOLO algorithm and Yolov5: The state-of-the-art object detection algorithm.

- 23.Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–9). IEEE.

- 24. Li, G., Ma, Z., & Wang, H. (2011, October). Image recognition of grape downy mildew and grape powdery mildew based on support vector machine.

In *International Conference on Computer and Computing Technologies in Agriculture* (pp. 151–162). Springer, Berlin, Heidelberg.

- 25. Xu, Y., Hu, M., Liu, H., Yang, H., Wang, H., Lu, S., Liang, T., Li, X., Xu, M., Li, L., & Li, H. (2021). A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis. *NPJ Digital Medicine*, 4(1), 48.

- 26.Luo, X., Li, J., Chen, M., Yang, X., & Li, X. (2021). Ophthalmic disease detection via deep learning with a novel mixture loss function. *IEEE Journal of Biomedical and Health Informatics*, 25(9), 3332–3339.

- 27. Patton, N., Aslam, T. M., MacGillivray, T., Deary, I. J., Dhillon, B., Eikelboom, R. H., Yogesan, K., & Constable, I. J. (2006). Retinal image analysis: Concepts, applications and potential. *Progress in Retinal and Eye Research*, 25(1), 99–127.

- 28.Carrasco, D. P., Rashwan, H. A., García, M. Á., & Puig, D. (2021). T-YOLO: Tiny vehicle detection based on YOLO and multi-scale convolutional neural networks. *IEEE Access*.

- 29.Kasper-Eulaers, M., Hahn, N., Berger, S., Sebulonsen, T., Myrland, Ø., & Kummervold, P. E. (2021). Detecting heavy goods vehicles in rest areas in winter conditions using YOLOv5. *Algorithms*, 14(4), 114.

- 30.Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software*, 95, 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

How to install LaTeX

Windows OS

TeXLive package - full version

1. Download the TeXLive ISO (2.2GB) from
   https://www.tug.org/texlive/
2. Download WinCDEmu (if you don't have a virtual drive) from
   http://wincdemu.sysprogs.org/download/
3. To install Windows CD Emulator follow the instructions at
   http://wincdemu.sysprogs.org/tutorials/install/
4. Right click the iso and mount it using the WinCDEmu as shown in
   http://wincdemu.sysprogs.org/tutorials/mount/
5. Open your virtual drive and run setup.pl

or

Basic MikTeX - TeX distribution

1. Download Basic-MiKTeX(32bit or 64bit) from
   http://miktex.org/download
2. Run the installer
3. To add a new package go to Start ¿¿ All Programs ¿¿ MikTex ¿¿ Maintenance (Admin) and choose Package Manager
4. Select or search for packages to install

TexStudio - TeX editor

1. Download TexStudio from
   http://texstudio.sourceforge.net/#downloads
2. Run the installer

Mac OS X

MacTeX - TEX distribution

1. Download the file from
   https://www.tug.org/mactex/
2. Extract and double click to run the installer. It does the entire configuration, sit back and relax.

TexStudio - TEX editor

1. Download TexStudio from
   http://texstudio.sourceforge.net/#downloads
2. Extract and Start

Unix/Linux

TeXLive - TEX distribution

Getting the distribution:

1. TexLive can be downloaded from
   http://www.tug.org/texlive/acquire-netinstall.html.
2. TexLive is provided by most operating system you can use (rpm,apt-get or yum) to get TexLive distributions

Installation

1. Mount the ISO file in the mnt directory

   ```
   mount -t iso9660 -o ro,loop,noauto /your/texlive####.iso /mnt
   ```

2. Install wget on your OS (use rpm, apt-get or yum install)
3. Run the installer script install-tl.

   ```
   cd /your/download/directory
   ./install-tl
   ```

4. Enter command 'i' for installation
5. Post-Installation configuration:
   http://www.tug.org/texlive/doc/texlive-en/texlive-en.html#x1-320003.4.1
6. Set the path for the directory of TexLive binaries in your .bashrc file

For 32bit OS

For Bourne-compatible shells such as bash, and using Intel x86 GNU/Linux and a default directory setup as an example, the file to edit might be

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/i386-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

For 64bit OS

```
edit $~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/x86_64-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

Fedora/RedHat/CentOS:

```
sudo yum install texlive
sudo yum install psutils
```

SUSE:

```
sudo zypper install texlive
```

Debian/Ubuntu:

```
sudo apt-get install texlive texlive-latex-extra
sudo apt-get install psutils
```

Overleaf: GitHub for LaTeX projects

This Project was developed using Overleaf(https://www.overleaf.com/), an online LaTeX editor that allows real-time collaboration and online compiling of projects to PDF format. In comparison to other LaTeX editors, Overleaf is a server-based application, which is accessed through a web browser.