International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V72I10P121
Volume 72 Issue 10, 141-147, October 2024


*Original Article*

# Comparative Analysis of Cataract Eye Disease Detection Using Yolov8 and Yolov10

Arkesha Shah

*Department of Information Technology, Charotar University of Science and Technology(CHARUSAT), Anand, Gujarat, India.*

*Corresponding Author : arkeshadesai@gmail.com*

Received: 07 September 2024 Revised: 09 October 2024 Accepted: 23 October 2024 Published: 31 October 2024

***Abstract*** - *Cataracts represent a major global health challenge, which affects millions of individuals and causes vision impairment. Which generally happens with older people. Early detection and timely diagnosis are important to mitigate the impact of cataracts and prevent early eye vision loss. Automatic detection of cataracts in the Eyes can greatly assist healthcare professionals in early diagnosis and management and the prevention of blindness, ultimately improving patients' health outcomes and reducing the burden on healthcare resources. This study proposes a cataract detection system using YOLOv8, a cutting-edge object detection model. By adapting YOLOv8 to the specific challenges of cataract detection, the aim is to develop a robust and efficient solution for automated cataract screening. The methodology involves training the YOLOv8 model on a comprehensive dataset of retinal images annotated with cataract labels. To assess the effectiveness of the proposed system, I evaluated its performance on a separate test set of retinal images, measuring key metrics such as precision, recall, and F1-score. The Comparative Performance Analysis of the YOLOV8 Model and YOLOV10 Model is also done. This evaluation aims to validate the system's ability to detect cataracts and its potential utility in clinical practice accurately*.

***Keywords*** - *Cataract Eye Disease, YOLOV8, YOLOV10, Deep Learning Model,Object detection.*

## 1. Introduction

In cataracts, the eye focal point becomes cloudy, which causes hazy vision and visual deficiency, especially among the elderly population [1]. Opportune discovery is imperative for viable treatment of cataracts, which makes strides in the vision of people and spares their lives since vision is exceptionally fundamental and vital for life [3]. Conventional strategies for cataract conclusion regularly depend on subjective appraisals by ophthalmologists, which can be time-consuming and may need consistency [5]. In later a long time, progressions in computer vision and profound learning have opened modern conceivable outcomes for robotized therapeutic picture examination, including the discovery of cataracts [7]. Convolutional Neural Network systems (CNNs) have illustrated surprising victories in different picture acknowledgement errands, advertising the potential to help healthcare experts precisely and proficiently diagnose cataracts [11]. You Merely See Once (YOLO) stands out for its real-time object-detection capabilities and tall exactness among the CNN models. YOLO forms pictures in a single pass through the organization, empowering the fast location of objects with exact bounding boxes and lesson probabilities. Recent progressions in YOLOv8 and YOLOv10 have improved their organized design and preparation techniques, making these adaptations especially successful for complex question location scenarios. This ponder presents a framework for recognizing cataracts that utilizes YOLOv8 and YOLOv10 calculations. The essential objective is to create a solid and effective strategy for computerized cataract screening. The framework is planned to back therapeutic experts in precisely and reliably recognizing cataracts through the examination of retinal pictures.

## 2. Literature Review

Hind Hadi Ali [1] et al., in 2022 their research introduced a Deep Convolution Neural Network (DCNN)-based automated cataract diagnosis approach. In their research, they collected fundus images of cataract disease from Kaggle. Then, those images are preprocessed for the DCNN Model so that it gives a better result. They used augmentation methods to avoid overfitting using the DCNN model, and they achieved 97% accuracy. Raghavendra Chaudhary[2] et al. In 2022, their research used Convolutional Neural Networks (CNN) to diagnose cataracts in the eyes; for that, they used the Digital Camera Image dataset, which is publicly available. Their proposed model consists of three convolutional layers, three pooling layers, a flattened layer, and two dense layers. They have used Optimizer as Adanm and Achieved 0.9924 Accuracy.Shuvam Chakraborty[3] et al. In this paper, they proposed a convolutional neural network model to detect cataracts in the eyes. They used VGG19 architecture, which has three layers. They used Prepossessing on images, and then they divided the dataset into training and testing. After testing the model on unseen data, they achieved 97.47%.Md. Sajjad Mahmud Khan[4] et al.In this paper, they proposed the VGG19 model, which is a convolutional neural network model to detect the cataract on the color fundus images. Their dataset contains 800 fundus images. From 800 images, they only utilized cataract and normal fundus images. Softmax and Relu activation functions were applied to the dataset to learn the complex pattern collected from Kaggle. They achieved 97.47% accuracy. Zaidatul Nisa Binti Abdul Basit[5] et al., in this paper, proposed a Convolutional Neural Network (CNN) model and Visual Geometry Group (VGG 19) for the detection of cataracts in the eyes. They applied Different Approaches to image augmentation, fine-tuning strategies, and various image resolutions for the detection. They did the different experiments. From that, they got an augmented VGG19 model that achieved an accuracy of 93.57%. Which is better performed than the unaugmented VGG19 model (92.72%) and the unaugmented CNN (89.14%). Kanwarpartap Singh Gill[6] et al. 2023 proposed the VGG19 model, which is capable of identifying Cataract Eyes and Normal Eyes.

The Proposed VGG-19 model achieved 90% Accuracy. Ahsan Abbas[7] et al., in 2023, introduced Deep learning models based on transfer learning. They are VGG19 and ResNet-50 Models; these models are used to detect cataracts in the Eyes. They used different performance metrics to check model performance. Their model achieved the highest accuracy of 98%. Deepak Kumar[8] et al. in This paper introduce a deep Learning Model that utilizes Convolutional Neural Network (CNN) models, namely VGG16 and ResNet50, and a Vision Transformer (ViT) based approach for the detection of cataracts in the eyes. They have introduced media noise filtering, which is implemented as median filtering; these techniques are employed as a preprocessing technique to reduce noise and improve overall image quality. They used data augmentation to avoid overfitting. ViT Model achieved the highest accuracy of 70%.Wajeeha Ahmed[9] et al. in this paper, introduced ResNet-18 and RestNet-50 Deep Learning Models. They used the Softmax Activation function in Pretrained Models. The RestNet-18 model achieved the highest accuracy of 98.9%.Masum Shah Junayed[10] et al.In this paper, They Proposed a deep neural network, namely CataractNet, for automatic cataract detection in fundus images. They used loss and activation functions to train the network. They used an optimizer like Adam to avoid the over-fitting problem and Applied Data Augmentation on images. CataractNet experimental results had the highest accuracy of 99.13%. Syeda Nabila Shirazi[11] et al., in this paper, proposed Convolutional Neural Networks, GoogleNet and AlexNet are proposed for detecting eye cataracts. They performed comparisons with all other models, and from that, the GoogleNet Model achieved the highest accuracy of 97%.Hima Bindu Koncha [21] et al. In this paper, they proposed that the VGG-19 model detects cataracts from the collected fundus images. They preprocessed the dataset before feeding the data into the model. The pre-trained model was further validated. After providing the Test Data, the model got the highest accuracy of 96.79%

## 3. Methodology

### *3.1. YOLO V8 Architecture*

YOLOv8 Profound Learning Demonstrate is exceptionally successful and supportive in tackling different industry issues, including mechanical technology, independent driving, and video surveillance. It has the capacity to distinguish objects for security and decision-making forms. The YOLOv8 Profound Learning Show is valuable to recognize and localize objects in pictures and recordings with productive accuracy. YOLOv8 Show is another cycle of the YOLO arrangement.

YOLO, a brief for You Simply See Once, was made in 2015.YOLO partitions the whole picture into a network, and each network is utilized to identify objects inside itself. YOLOv8 design advances past YOLO models, which have outflanked prior forms by including diverse adjustments such as spatial consideration, highlight combination, and setting accumulation modules. YOLOv8 is the YOLO form that was discharged in January 2023. The design of YOLOv8 depends on the prior forms of YOLO calculations. YOLOv8 is comprised of basically two parts: the backbone and the head.

#### *3.1.1. Powerful Backbone*

YOLOV8 employs a pre-trained convolutional neural arrange profound learning demonstration such as Darknet or Efficient-53Det, which is utilized to extricate significant highlights from input pictures.

#### *3.1.2. Refinement and Enhancement*

YOLOv8 employs the diverse Procedures,

- Spatial Attention Module (SAM): This Strategy is Outlined to center on the foremost extricated highlights, which makes strides in discovering little or impeded objects.
- Path Aggregation Network (PAN): This combines diverse highlights from distinctive arrange layers and combines low-level points of interest with high-level data for a nitty gritty understanding of the picture.

#### *3.1.3. The Prediction Layers*

- Grid Division: The whole picture is separated into a network of cells, such as a 16×16 framework. Each cell could detect potential objects. Within each cell, the model can predict:
- Bounding Boxes: it will Predict several bounding boxes of different sizes and aspect ratios,
- Confidence Scores: it will predict a confidence score for each bounding box; it believes that the bounding box containing an object and the correct class Label

- Class Probabilities: Predicts the probability of each object and its corresponding class

*3.1.4. Loss Functions*

During the training period, loss functions like Intersection over Union (IoU) IoU Loss will ensure correct surroundings around the actual objects.

Classification Loss: If the error is minimized, it will be helpful to predict class probabilities, which leads to more accurate object classifications.

*3.1.5. Output and Post-Processing*

Non-Maxima Suppression (NMS) can be applied to refine detections by removing redundant bounding boxes.

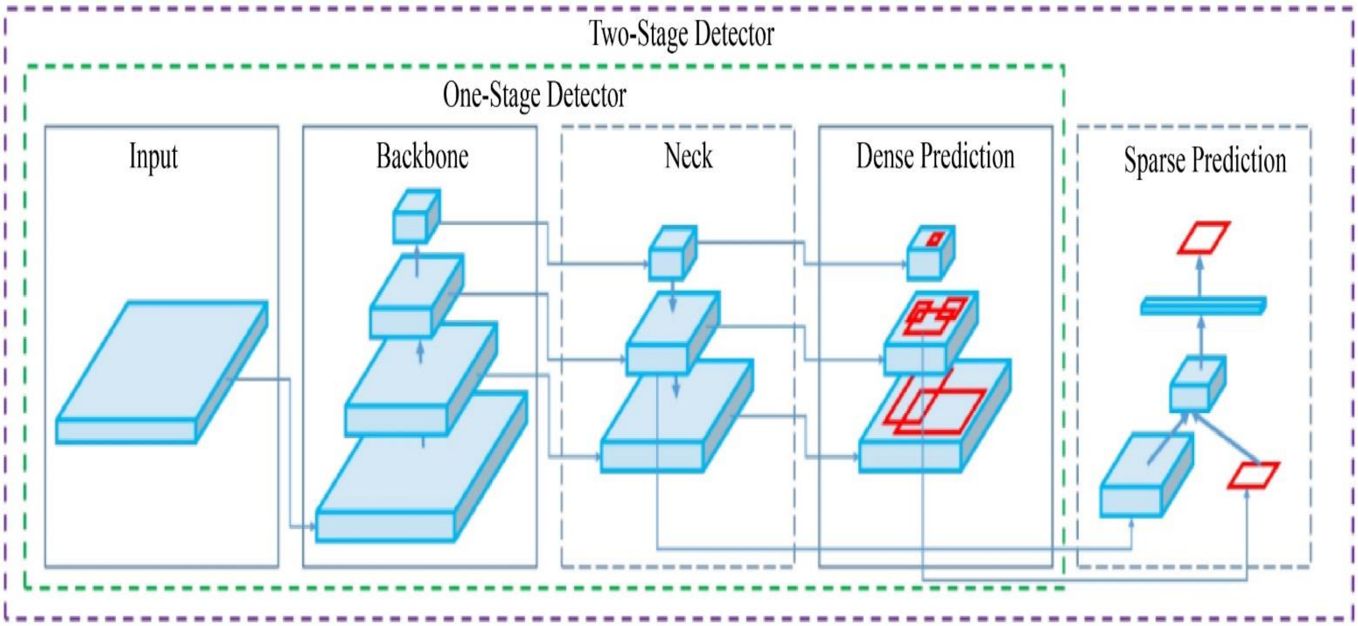


**Fig. 1 Working of the YOLO model**

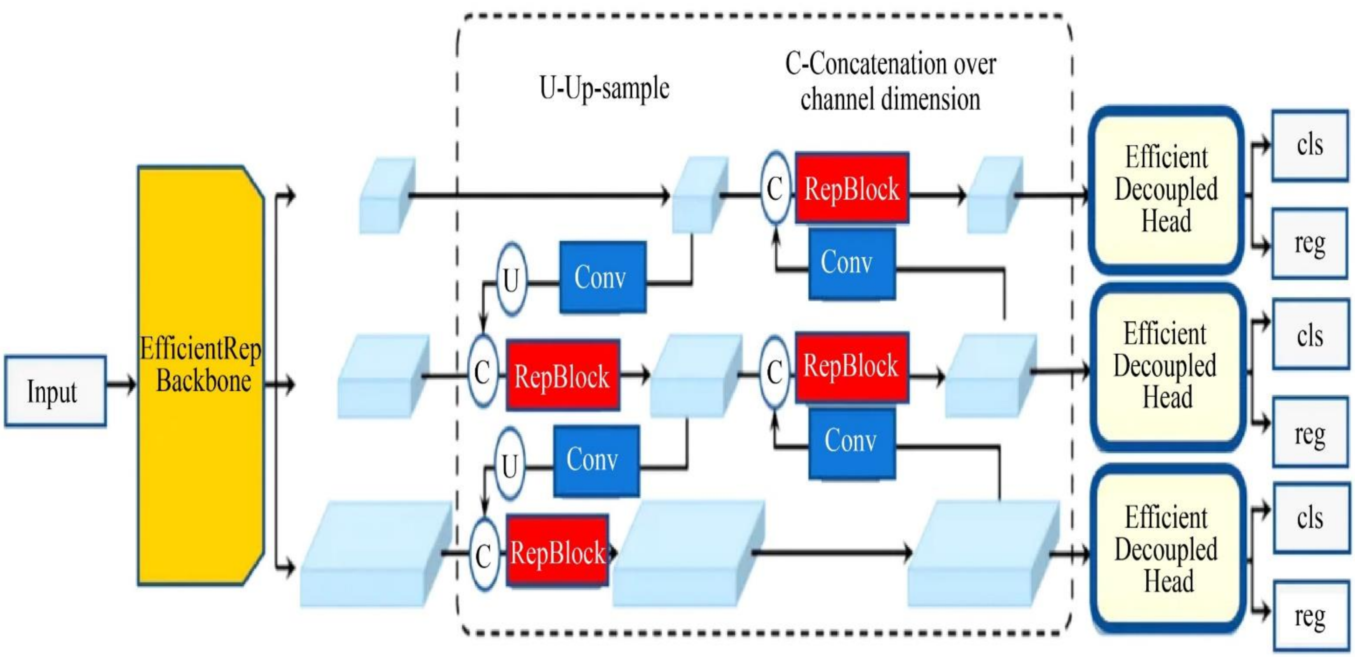


**Fig. 2 YOLOV8 Architecture**

### *3.2. YOLOV10 Architecture*

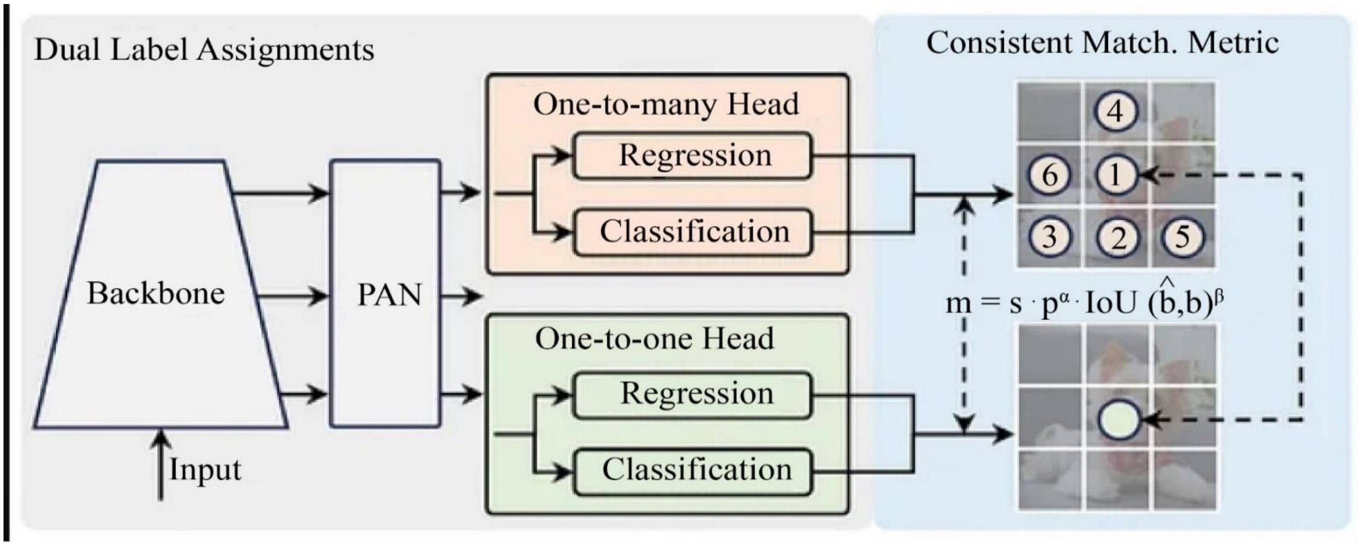


**Fig. 3 YOLOV10 Architecture**

YOLOv10 presented an NMS-free object Detection Technique. The engineering of YOLOV10 is based on the past adaptation of YOLO models. The Architecture consists of four Components. Backbone, Neck, One to Many head and one to one head. The Key Highlights are the holistic model Design, which contains lightweight classification heads and spatial Chanel decoupled-down sampling.

### *3.3. Dataset Preparation*

The Cataract Eye disease images were collected from Roboflow Universe, which Contains Open Source Dataset. The images were annotated using the Roboflow Annotation tool. It does not need any installation on the PC. The annotation was done by drawing rectangle-bounding boxes around the affected eye area; then, the boxes were labeled with classes and exported using YOLO format. The label file in YOLO format has 5 values for each bounding box. The first value contains the class of the bounding box, followed by the x_center, y_center, width, and height of the bounding box, where x_center and y_center are the center point of the bounding boxes. All these values are normalized in between the range of 0 to 1. The collected dataset Contains a total of 1015 images from 711 images used for Training,203 images for validation, and 101 images for testing.

### *3.4. Training the Custom YOLOv8-YOLOV10 Model*

The model is trained on Google Colab, a Jupyter Notebook environment. Anyone can write and run Python code in the browser. For that, Internet Connectivity is needed. It is easy and efficient to use because all the necessary pre-installed libraries required for object detection tasks are easily imported. It also provides free access to computing resources, such as GPUs, for some period of time, which makes it much faster. The model parameters for training input images are the size of 640×640×3, batch size of 64 and number of epochs of 100.

### *3.5. Metrics for Evaluation*

In evaluating the performance of the proposed YOLOv8 model, several metrics are used, Such as Precision, Recall, F1 score, Mean average precision (mAP) and Confusion matrix. The following terms are used to check the model Performance:

- True positive (TP): Eyes with Actual Cataracts, and Model correctly predicted as Eyes have cataracts.
- False positive (FP): Eyes which do not have cataracts but are falsely predicted as cataracts.
- False negative (FN): Eyes that have cataracts but are falsely predicted as not cataracts.
- True negative (TN): Eyes that do not have cataract and correctly predicted as not cataract.
- Precision is a metric that checks how nearly the predictions are from the actual truth.

$$Precision + True\ Positive / True\ Positive + False\ Positive \tag{1}$$

Recall is a matrix used to check from all real Positive cases how many cases the model has Predicted Positives.

$$\text{Recall} = True\ Positive\ /\ True\ Positive + \text{False Negative} \tag{2}$$

The F1 Score is used to determine the testing accuracy of the model, which falls between 0 and 1. The harmonic mean of recall and precision is recognized by the F1 Score.

$$\text{Accuracy} = True\ Positive + \text{True Negative / Total} \tag{3}$$

Mean average precision (mAP): It is used for determining the model's accuracy.mAP is the average value of area under the precision-recall curve.

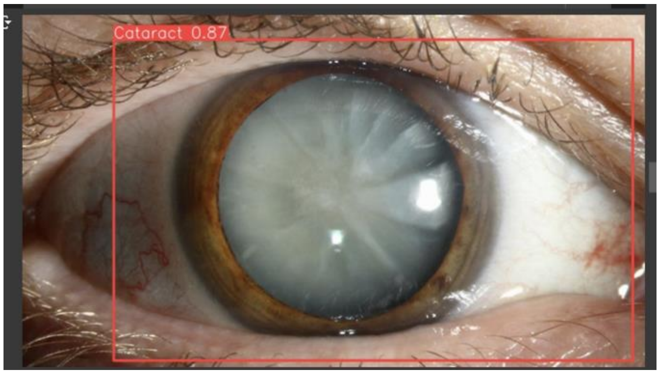


**Fig. 4 Results of model detection on test images**

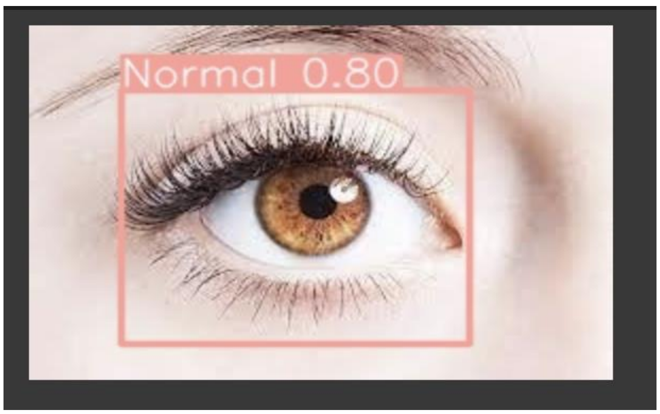


**Fig. 5 Results of model detection on test images**

## 4. Results and Discussion

The dataset is divided into Training, Testing, and Validation, defined as 70% of the data for training, 20% for testing, and 10% for validation. The dataset contained 1015 images from 711 images used for Training,203 images for validation, and 101 images for testing. The YOLOv8 model is trained using the Cataract Eye Disease dataset for the early detection and Prevention of Eye Blindness. The training speed was 0.4 ms. The model is prepared for over 100 epochs. After Training, the Yolov8 model had 295 layers, which contained 25857478 parameters. In a deep neural network, the term "loss" is the error that occurs during the prediction. Training loss is used to check how the model is applied to the data used to train it.

In a deep neural network, the term "loss" denotes the error during the prediction. The training loss is used to check how the model is applied to the data used to train it. The validation loss determines the model's performance on unseen data. Box loss is used to recognize the center of the object and correctly identify the bounding box that encloses the object. The class loss determines how well the model can predict the actual class of a given object. The model yielded good results after training over 100 epochs in accordance with precision, recall, and mAP@.5. Mean Average Precision (mAP) is used for analyzing the performance of object detection models. A score of 0.5, the Average Precision for balanced data, and the model's precision at an IoU threshold of 0.5, indicates how well it can detect objects correctly; a score of 1.0 represents a perfect model. YOLOv8 deep learning model is popular for its high inference speed, and in this work, it took 10.8 ms for inference per image. The model achieved the highest accuracy (95.7 %).

**Table 1. Results model YOLOV8**

| YOLO Model | mAP@.5 | Precision | Recall |
|---|---|---|---|
| YOLOV8 | 95.7% | 0.92 | 0.93 |

## 5. Comparative Result Analysis

Cataract Eye disease detection was done using the YoloV8 and YOLOV10 Models. In this Research, The Dataset was fed to the Yolov8 Model, which Had 295 layers and contained 25857478 parameters. The model took a speed of 0.4 ms to preprocess, 10.8 ms for inference and 7.3 ms for postprocess; then, after the Same Dataset was fed to the Yolov10 Model, which had the 369 layers, the 16452700-parameter model took a speed of 0.3 ms to preprocess,25.1ms for inference and 1.0ms for post process per image. The results show that the YOLOV10 Model Comparatively Performed Better than the YOLOV8 Deep Learning Model. Post Process time of Yolov10 is better than the YOLOV8 Deep Learning Model because it is considered a NMS Free Architecture. The preprocess speed of Yolov10 is better as compared to the YOLOV8 Deep Learning Model, and the inference speed of the YOLOV8 Deep Learning Model is better compared to the Yolov10 Deep Learning Model.

**Table 2. Analysis of results between YOLOV8 AND YOLOV10**

| YOLO Model | Epochs | Batch Size | mAP@.5 | Precision | Recall |
|---|---|---|---|---|---|
| YOLOV8 | 100 | 16 | 95.7% | 0.92 | 0.93 |
| YOLOV10 | 100 | 16 | 98.3% | 0.96 | 0.97 |

## 6. Conclusion

The proposed system detects cataracts from the eye's image. This task was done using the YOLOV8 and YOLOV10 Deep Learning Model. The Dataset Contains Two Classes, which are label as cataract and normal. The model achieved 98.9% accuracy after successfully detecting cataracts from the eyes. The model preprocessed the data at the speed of 0.4 ms, which was done on Google Colab, which is Cloud PlatForm, which made the Process Easy.

The Same Data fed into the YoloV10 Deep Learning Model, which achieved an accuracy of 98.3%. The Comparative Analysis between the YOLOV8 and YOLOV10 Deep Learning Models was done in this research. From that, the YOLOV10 model outperformed the YOLOV8 Model in terms of accuracy, better efficiency, and reduced inference latency. The proposed system has the potential to assist healthcare professionals in early diagnosis and the prevention of blindness, thus contributing to improved patient health outcomes and reducing the efforts of healthcare resources.

## References


[1] Hind Hadi Ali, Ali Yakoob Al-Sultan, and Enas Hamood Al-Saadi, "Cataract Disease Detection Used Deep Convolution Neural Network," *2022 5th International Conference on Engineering Technology and its Applications*, Al-Najaf, Iraq, pp. 102-108, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Raghavendra Chaudhary, and Arun Kumar, "Cataract Detection Using Deep Learning Model on Digital Camera Images," *2022 IEEE International Conference on Cybernetics and Computational Intelligence*, Malang, Indonesia, pp. 489-493, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Shuvam Chakraborty, and Susovan Jana, "Early Prediction of Cataract using Convolutional Neural Network," *2023 IEEE Devices for Integrated Circuit*, Kalyani, India, pp. 446-450, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Md. Sajjad Mahmud Khan et al., "Cataract Detection Using Convolutional Neural Network with VGG-19 Model," *2021 IEEE World AI IoT Congress*, Seattle, WA, USA, pp. 209-212, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Zaidatul Nisa'Binti Abdul Basit, and Yoshihiro Mitani, "Cataract Disease Detection Based on Small Fundus Images Dataset Using CNNs," *2024 2nd International Conference on Computer Graphics and Image Processing*, Kyoto, Japan, pp. 113-117, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Kanwarpartap Singh Gill, Vatsala Anand, and Rupesh Gupta, "Cataract Detection Using optimized VGG19 Model by Transfer Learning perspective and its Social Benefits," *2023 Second International Conference on Augmented Intelligence and Sustainable Systems*, Trichy, India, pp. 593-596, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Ahsan Abbas et al., "A Transfer Learning Based Detection and Grading of Cataract using Fundus Images," *2023 25th International Multitopic Conference*, Lahore, Pakistan, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Deepak Kumar et al., "Cataract Disease Identification Using Transformer and Convolution Neural Network: A Novel Framework," *2023 3rd International Conference on Technological Advancements in Computational Sciences*, Tashkent, Uzbekistan, pp. 1230-1235, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Wajeeha Ahmed et al., "Automatic Diagnosis of Cataract and Myopia through Fundus Images," *2023 International Conference on Business Analytics for Technology and Security*, Dubai, United Arab Emirates, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Masum Shah Junayed et al., "CataractNet: An Automated Cataract Detection System Using Deep Learning for Fundus Images," *IEEE Access*, vol. 9, pp. 128799-128808, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Syeda Nabila Shirazi et al., "Computer-Aided Diagnosis of Cataract Disease through Retinal Images," *2022 International Conference on IT and Industrial Technologies*, Chiniot, Pakistan, pp. 1-7, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] D. Shamia, Shajin Prince, and D. Bini, "An Online Platform for Early Eye Disease Detection using Deep Convolutional Neural Networks," *2022 6th International Conference on Devices, Circuits and Systems*, Coimbatore, India, pp. 388-392, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Natalia Sokolova et al., "Pixel-Based Iris and Pupil Segmentation in Cataract Surgery Videos Using Mask R-CNN," *2020 IEEE 17th International Symposium on Biomedical Imaging Workshops*, Iowa City, IA, USA, pp. 1-4, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Ahmad Bondan Triyadi, Alhadi Bustamam, and Prasnurzaki Anki, "Deep Learning in Image Classification Using VGG-19 and Residual Networks for Cataract Detection," *2022 2nd International Conference on Information Technology and Education*, Malang, Indonesia, pp. 293-297, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Renato R. Maaliw et al., "Cataract Detection and Grading Using Ensemble Neural Networks and Transfer Learning," *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference*, Vancouver, BC, Canada, pp. 74-81, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] P. Nageswari et al., "Automatic Detection and Classification of Diabetic Retinopathy Using Modified UNET," *2023 Third International Conference on Artificial Intelligence and Smart Energy*, Coimbatore, India, pp. 1468-1471, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Riya Sharma et al., "Modified EfficientNetB3 Deep Learning Model to Classify Colour Fundus Images of Eye Diseases," *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications*, Hamburg, Germany, pp. 632-638, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Urmila Pilania et al., "An Optimized Hybrid Approach to Detect Cataract," *2022 IEEE Global Conference on Computing, Power and Communication Technologies*, New Delhi, India, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Pankaj Kumar, Sudhir Bhandari, and Vishal Dutt, "Pre-Trained Deep Learning-Based Approaches for Eye Disease Detection," *2023 International Conference on Circuit Power and Computing Technologies*, Kollam, India, pp. 1286-1290, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] K.M. Mayalekshmi, Abhishek Ranjan, and Rajendra Machavaram, "In-field Chilli Crop Disease Detection Using YOLOv5 Deep Learning Technique," *2023 IEEE 8th International Conference for Convergence in Technology*, Lonavla, India, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Hima Bindu Koncha et al., "Early Detection and Prediction of Cataract Using Deep Learning," *2023 IEEE 8th International Conference for Convergence in Technology*, Lonavla, India, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]