Real-Time Cataract Diagnosis with GhostYOLO: A GhostConv-enhanced YOLO Model

Lahari Puchakayala Lokesh
Department of Electronics and Communication Engineering, SRM University, Andhra Pradesh, India
lahari_p@srmap.edu.in (corresponding author)

Rahul Gowtham Poola
Department of Computing Technologies, SRM Institute of Science and Technology, Tamil Nadu, India
rahulgowtham_poola@srmap.edu.in

Leela Prasad Gorrepati
Camelot Integrated Solutions INC, Virginia, USA
leelaprasad.gorrepati@gmail.com

Siva Sankar Yellampalli
Department of Electronics and Communication Engineering, SRM University, Andhra Pradesh, India
sivasankar.y@srmap.edu.in

Received: 1 March 2025 | Revised: 23 March 2025 | Accepted: 4 April 2025

 DOI: https://doi.org/10.48084/etasr.10760

ABSTRACT

This study presents GhostYOLO, an enhanced YOLO-based model for cataract detection that incorporates GhostConv layers to offer greater accuracy, faster processing, and less memory consumption for real-time diagnosis. Initially, the performance of YOLO models, namely YOLOv5, YOLOv6, YOLOv7, and YOLOv8, was evaluated using 788 normal and 920 cataract images, with YOLOv8n emerging as the best standard model with excellent precision, speed, and efficiency. GhostYOLO models were developed to further improve speed and accuracy. GhostYoloV8n obtained the highest accuracy, speed, and lowest memory usage, while GhostYoLoV7-tiny also performed well. Incorporating GhostConv layers substantially improved cataract detection, increasing efficiency and real-time usage. Real-time tests using a Jetson Nano board demonstrated its efficiency, with 33.5 FPS in live detection, simplifying diagnosis. GhostYoloV8n, with only 1.6 million parameters, is a small but powerful cataract detection tool that allows for faster and more precise medical intervention. This study highlights the benefits of including GhostConv layers in YOLO models, making cataract diagnosis more accurate, efficient, and scalable for clinical applications.

Keywords-cataract detection; YOLO models; GhostYOLO; GhostConv layers; real-time diagnosis; Jetson nano

I. INTRODUCTION

Cataracts [1, 2], which are defined by the clouding of the eye's natural lens, are one of the main causes of blindness globally, disproportionately affecting the elderly [3, 4]. This disorder occurs when proteins in the lens clump together, obstructing light flow and resulting in blurred or distorted vision. Common symptoms include hazy vision, increased sensitivity to light, and difficulty seeing in low-light situations [5]. These symptoms have a significant influence on daily tasks such as reading, driving, and recognizing faces, lowering the quality of life of those affected [6]. Cataracts become more common as people age, making them an important public health concern in older populations. The World Health Organization (WHO) has identified cataracts as a major cause of global eyesight loss [7-9]. Cataracts can often be treated surgically, with the clouded lens removed and replaced with an artificial intraocular lens. Early diagnosis is critical to prevent serious vision damage, as timely surgical surgery can significantly improve visual acuity and patient satisfaction [10, 11]. Despite the availability of viable therapies, millions of people remain blind, mainly in low- and middle-income countries, due to a lack of diagnostic and surgical services. This demonstrates the critical need for cost-effective and scalable solutions to overcome the cataract care gap.

Recent advances in Artificial Intelligence (AI) and computer vision have altered the landscape of cataract detection and management [12, 13]. AI-powered diagnostic techniques, trained on large clinical image datasets, allow the early and precise detection of cataracts, frequently outperforming traditional diagnostic approaches in terms of speed and precision [14, 15]. These technologies not only provide consistent and objective assessments but can also be integrated into portable devices and used in rural or underserved areas. These improvements enhance the efficiency of cataract screening, expand access to care, and have the potential to drastically reduce the global burden of cataracts and improve vision outcomes for millions of people around the world. Machine learning algorithms have transformed the field of ophthalmology by allowing for more accurate and efficient detection of cataracts. Early diagnosis with these automated devices is critical because it allows for quick intervention, reducing the risk of complications and increasing long-term patient outcomes. Furthermore, these AI-powered solutions alleviate the pressure on healthcare workers, allowing them to focus on complex cases that require their expertise.

YOLO (You Only Look Once) object identification models [16] have emerged as important tools in medical imaging, particularly for real-time cataract detection. Known for their unusual mix of speed and precision, YOLO models perform object recognition in a single pass through the neural network, making them very efficient for real-time applications. This skill is crucial in cataract screening, as rapid analysis allows timely decision-making in clinical workflows. The versatility of YOLO models in numerous imaging modalities, including slit-lamp photography, fundus imaging, and anterior segment scans, strengthens their importance in a wide range of healthcare settings. Their implementation not only improves diagnostic speed but also reduces total screening costs by streamlining procedures. Despite these benefits, traditional YOLO models have drawbacks when used in resource-constrained devices such as portable imaging equipment and wearable sensors. These devices, which are frequently utilized in rural or underserved areas, necessitate models that are optimized for processing speed and fewer computing resources while preserving accuracy. To solve this issue, improvements in hardware-accelerated solutions and lightweight model architectures have been proposed. By combining efficient YOLO variations with portable diagnostic instruments, healthcare providers can offer high-quality cataract screening to groups who have previously lacked access to specialized care. This combination of AI-powered diagnostics and accessible hardware solutions is a huge step toward egalitarian healthcare, improving vision care results on a worldwide scale.

Various YOLO-based approaches have been explored for cataract detection. In [17], eye lens recordings were used for cataract detection using YOLOv3 for precise lens localization and color space conversion for cataract classification. The dataset included 76 eyes from 38 patients, totaling 1,520 images divided into training, validation, and testing (7:2:1 ratio). This method achieved 94% accuracy and 93.88 % F1 score on clinical test data. In [18], the focus was on Diabetic Retinopathy (DR) detection in the Saudi Arabian population, where existing systems suffer from high computational costs and dataset imbalances. A low-complexity deep learning model was developed using YOLOv7 for feature extraction and the Quantum Marine Predator Algorithm (QMPA) for feature selection. The MobileNet V3 model achieved 98.0% and 98.4% accuracy on the APTOS and EyePacs datasets, respectively, requiring fewer computational resources and making it suitable for mobile applications in healthcare. In [19], early cataract detection was performed using smartphone-based slit-lamp images. YOLOv3 localized the nuclear region of the ocular lens, and ShuffleNet combined with an SVM classifier rated cataract severity. The model achieved 93.5% accuracy, 92.3% F1 score, and 95.4% Kappa. The system was designed for rapid evaluation (29 ms per image), offering a low-cost solution for cataract screening in resource-limited settings. In [20], YOLOv8 was used for cataract detection, achieving 96.1% accuracy and outperforming other models, such as VGGNet, AlexNet, and ResNet50, in both speed and accuracy, making it suitable for real-time applications. In [21], a system combined a camera module and a Raspberry Pi 4B with YOLOv5 to detect eye conditions in cats [22], including glaucoma, cataracts, and cherry eyes, achieving 90% accuracy across all disease classes and providing insights into improving animal health diagnostics.

All the studies on cataract diagnosis using YOLO models achieved great accuracy but sometimes battled with computational efficiency and applicability for real-time deployment, particularly on resource-limited devices. Many versions require substantial computing power, which limits their usefulness in portable or low-cost medical installations. GhostYOLO bridges this gap by using GhostConv layers, which improve feature extraction while lowering computational cost.

II. METHODOLOGY

This research introduces GhostYOLO models, an advanced version of the YOLO framework designed to improve cataract detection. By integrating GhostConv layers, GhostYOLO models achieve higher precision, faster speeds, and lower memory usage. These models address current limitations in cataract screening with improved real-time performance. The study evaluates various YOLO models, such as YOLOv5, YOLOv6, YOLOv7, and YOLOv8, using a dataset of 788 normal and 920 cataract images, finding YOLOv8n to be the best standard model. GhostYOLO variants were developed and rigorously tested to improve cataract detection accuracy. GhostYOLO models greatly improve cataract detection, helping doctors diagnose and treat cataracts more effectively. The compact and efficient design of the proposed system makes it suitable for real-time use, improving patient outcomes with rapid and accurate detection. As shown in Figure 1, the proposed framework begins with a dataset comprising 1,708 fundus images (788 normal and 920 cataracts). The dataset is split into training (80%), testing (10%), and validation (10%) subsets. The proposed model is then trained and validated using this dataset. Then, the model is deployed on a Jetson Nano board, equipped with a Lenovo CMOS camera. The Jetson Nano processes real-time input, and the output, displayed on an attached desktop, indicates whether the image represents a normal or cataract condition, along with a confidence score.

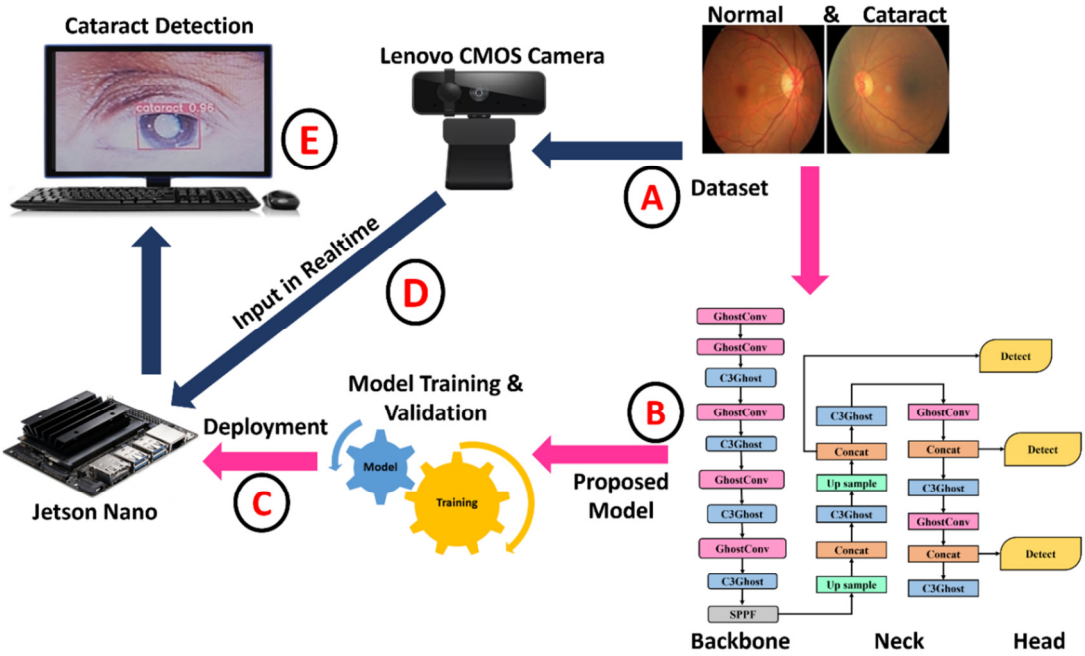

Fig. 1. Overview of the proposed approach.

A. *Dataset*

This study used a dataset from the Roboflow Dataset Repository [23] that consists of 427 images. These were divided into 343 images for the training set, 42 for the validation set, and the remaining for testing. Figure 2 provides training image samples. To ensure diversity, the dataset includes images with varying models, angles, resolutions, sizes, and orientations. However, the dataset initially exhibited class imbalance, with more images in the normal than in the cataract class. To mitigate this issue and enhance model performance, image augmentation was employed, expanding the dataset from 427 to 1,708 images using the Roboflow augmentation tool. Techniques such as random horizontal flipping, rotation, and horizontal and vertical shearing (8° and 9°, respectively) were applied. After augmentation, the dataset comprised 920 cataract images and 788 normal images. The dataset was then split into training (80%), validation (10%), and testing (10%) subsets.

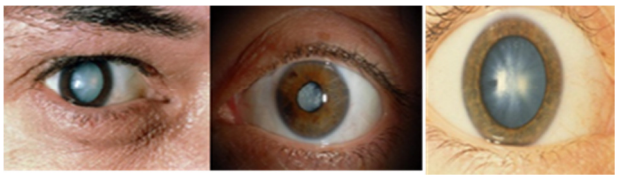

Fig. 2. Sample training images of cataract eyes.

B. *Proposed Model*

Object detection models have made significant advances in the past decade, with the evolution of the YOLO [24] series being a key development. YOLO reframed object detection as a single-pass regression problem, allowing for simultaneous prediction of bounding boxes and class probabilities, improving both speed and accuracy. Since its inception, YOLO has evolved rapidly. YOLOv2 [25] added batch normalization to improve convergence and reduce overfitting, increasing mAP by 2%. YOLOv3 [26], featured a deeper architecture, multiscale predictions, and residual connections, further enhancing accuracy while maintaining real-time speed. YOLOv4 [27] was optimized for both speed and accuracy. YOLOv5 [28] became popular for its efficiency and user-friendly design. Later versions, such as YOLOv6 [29], designed for edge devices, and YOLOv7 [30, 31], focused on optimizing real-time detection with novel strategies, such as compound model scaling. YOLOv8 [32], supports tasks such as object detection, segmentation, and classification. YOLOv8n, the nano version, introduces the c2f module and a decoupling head technique, resulting in faster detection speeds and higher accuracy.

Similarly to YOLOv8, the proposed model, shown in Figure 3, consists of three components: the backbone, neck, and head. The backbone is responsible for feature extraction from input images, the neck acts as a transition layer, and the head generates predictions based on refined features. In the proposed architecture, the backbone includes convolutional layers, GhostConv layers, C3Ghost blocks, and SPPF (Spatial Pyramid Pooling Fast). Convolutional layers extract spatial features by applying filters to the input image and detecting patterns.

- The GhostConv module splits convolution operations into two parts: the first performs traditional convolution and the second uses simple operations on feature maps, reducing computation and increasing efficiency.
- C3Ghost involves the use of ghost convolutional layers within a cross-stage structure.
- SPPF captures multi-scale information by applying max-pooling to grids of different sizes, enhancing feature extraction.

- The neck consists of 12 layers that combine FPN (Feature Pyramid Network) and PAN (Path Aggregation Network). This combination transfers spatial features across layers, improving object detection at different scales and enabling precise target identification regardless of size.
- The head then identifies object categories and outlines them with bounding boxes based on the extracted features.

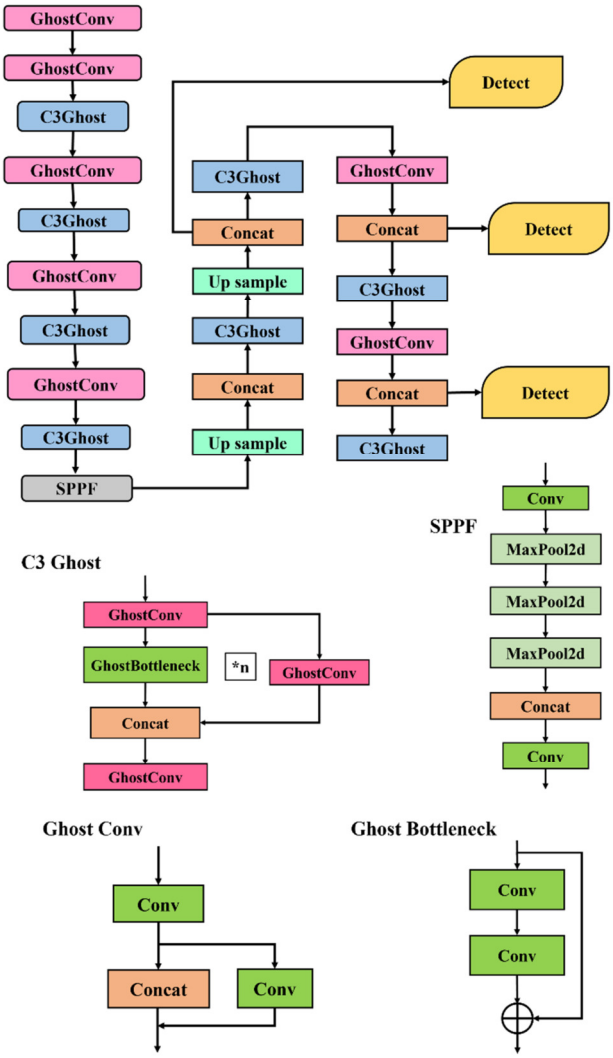

Fig. 3. Proposed model architecture and its blocks.

The GhostConv block begins by extracting basic features using the first convolution and then concatenates them with features from a secondary pathway. A second convolution refines the feature set, enhancing the model's classification ability. The C3Ghost module includes three GhostConv layers and multiple bottleneck modules, using lightweight convolution techniques to generate "ghost" feature maps. This approach reduces computational cost while maintaining effective feature extraction. The SPPF block further improves detection by optimizing multi-scale pooling, ensuring efficient processing and higher precision. Standard YOLO models face significant computational and memory demands, making real-time processing problematic on resource-constrained platforms. GhostYOLO addresses these limits by including lightweight components such as GhostConv layers, which reduce complexity while preserving speed, making it perfect for real-time applications.

III. RESULTS AND DISCUSSION

All tests were carried out on a system running Ubuntu 20.04 with Python 3.8, PyTorch 1.10.0-GPU, CUDA 11.3, and CUDNN 8.2.2, ensuring compatibility and leveraging GPU acceleration for efficient training and testing. The models were trained using pre-trained weights and maintaining consistent configurations across all of them. The training parameters included 300 epochs, a batch size of 8, 4 workers, a momentum of 0.95, and a learning rate of 0.001. The experimental setup utilized the Nvidia Jetson Nano board, a compact edge computing platform to develop the cataract detection and classification system. Using Jetson Nano hardware acceleration, the system was designed for real-time inference tasks.

A. Evaluation Metrics

Some commonly used metrics for evaluating the effectiveness of detection models include Precision (P), Recall (R), mean Average Precision (mAP), and F1 score (F1). The formulas for P, R, F1, and mAP are as follows:

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

$$mAP = \frac{\sum_{n=1}^{M} APn}{M} \quad (3)$$

$$\mathrm{F1} = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

These metrics are calculated based on True Positives (*TP*), False Positives (*FP*), and False Negatives (*FN*). *TP* represents the number of objects that have been correctly identified, while *FP* denotes the number of incorrectly detected objects. *FN* refers to the number of incorrectly undetected objects. The sum of *TP* and *FP* corresponds to the overall number of detected objects, and the sum of *TP* and *FN* represents the total number of actual objects. The F1 score is the harmonic mean of precision and recall. Precision refers to the ratio of *TP* to the total number of identified objects, while recall refers to the ratio of *TP* to the total number of actual objects. *N* represents the total number of classes in the test sample, and the Precision-Recall (*P-R*) curve area is known as average precision (AP). The average number of APs for different classes is called *mAP*.

B. Simulation Results

The YOLO family has advanced significantly from YOLOv5 to YOLOv8, enhancing both speed and accuracy in real-time object detection. YOLOv5 introduced modular variants (nano, small, medium, large, and extra-large) for different performance needs. YOLOv6 optimized training and inference for diverse applications. YOLOv7 added more powerful layers and models, including lightweight versions

such as YOLOv7-tiny. YOLOv8 continues this trend with variants ranging from nano to extra-large, offering state-of-the-art feature extraction and performance improvements for various real-time applications.

Simulation research compared the performance of various YOLO models (YOLOv5, YOLOv6, YOLOv7, and YOLOv8) on the dataset of cataract images, and Table I shows the results. Each model variant, nano (n), small (s), medium (m), large (l), and extra-large (x), is compared based on the number of layers, trainable parameters, GFLOPS (giga floating-point operations per second), precision, recall, mAP50 (mean average precision at 50% IoU), FPS (frames per second) speed, and memory usage during inference (in MB). YOLOv5n achieved fast speed (153.6 FPS), high precision (0.938), and low memory usage (6.3MB), while larger models such as YOLOv5x and YOLOv8x offered higher accuracy but slower speeds. YOLOv8n stands out with the highest precision (0.962), quick speed (175.3 FPS), and low memory consumption, making it the best standard model for cataract detection across all YOLO frameworks tested.

Table II compares cataract diagnosis models, specifically YOLO nano models and GhostConv-enhanced YOLO nano models. The comparison covers factors such as the number of layers, parameters, GFLOPS, precision, recall, mAP50, FPS performance, and memory usage in MB. All YOLO nano models with the GhostConv framework for cataract detection showed superior performance compared to standard YOLO models. The enhanced GhostConv models, which have more layers and parameters, also demonstrated improved computational efficiency with higher GFLOPS. Among these, GhostYoloV8n had 491 layers and 1.6 million parameters, achieving the highest recall (0.919) and precision (0.981), with a mAP50 of 0.954 at 191.6 FPS while utilizing only 3.7 MB of memory.

TABLE I. PERFORMANCE OF VARIOUS YOLO MODELS IN THE DETECTION OF CATARACT

Model	Layers	Parameters	GFLOPS	Precision	Recall	*mAP50*	Speed (FPS)	Memory (MB)
YoloV5n	262	2508854	7.2	0.938	0.856	0.932	153.6	6.3
YoloV5s	262	9122966	24.0	0.917	0.86	0.925	125.3	18.5
YoloV5m	339	25066294	64.4	0.908	0.881	0.918	85.2	50.5
YoloV5l	416	53164886	135.3	0.89	0.929	0.933	55.0	106.8
YoloV5x	493	97201334	246.9	0.804	0.887	0.893	35.5	194.9
YoloV6n	195	4238342	11.94	0.931	0.922	0.925	140.3	8.7
YoloV6s	195	16306230	44.2	0.899	0.905	0.91	155.7	32.8
YoloV6m	273	51997798	161.6	0.878	0.811	0.856	105.1	104.3
YoloV6l	351	110896278	391.9	0.578	0.672	0.66	75.6	222.2
YoloV7-tiny	255	70819527	13.2	0.892	0.897	0.914	167.4	7.2
YoloV7x	407	37200095	105.1	0.789	0.89	0.929	105.3	65.3
YoloV7	459	6016735	188.9	0.88	0.871	0.839	45.8	151.8
YoloV8n	225	3011238	8.2	0.962	0.915	0.948	175.3	5.3
YoloV8s	225	11136374	28.6	0.941	0.906	0.947	135.8	22.5
YoloV8m	295	25857478	79.1	0.931	0.863	0.931	95.0	52.0
YoloV8l	365	43631382	165.4	0.92	0.944	0.943	65.3	87.6
Yolov8x	365	68154534	258.1	0.896	0.904	0.946	40.2	136.7

TABLE II. PERFORMANCE OF VARIOUS YOLO NANO MODELS & NANO WITH GHOSTCONV IN THE DETECTION OF CATARACT

Algorithm	Model	Layers	Parameters	GFLOPS	Precision	Recall	*mAP50*	Speed (FPS)	Memory (MB)
Yolo nano frameworks for cataract detection	YoloV5n	262	2508854	7.2	0.938	0.856	0.932	153.6	6.3
	YoloV6n	195	4238342	11.94	0.931	0.922	0.925	140.3	8.7
	YoloV7-tiny	255	70819527	13.2	0.892	0.897	0.914	167.4	7.2
	YoloV8n	**225**	**3011238**	**8.2**	**0.962**	**0.915**	**0.948**	**175.3**	**5.3**
Yolo nano with GhostConv frameworks for enhanced cataract detection	GhostYoloV5n	529	2,200,340	5.1	0.971	0.894	0.946	169.2	5.8
	GhostYoloV6n	613	3,891,150	23.5	0.946	0.897	0.953	154.7	32.5
	GhostYoloV7-tiny	759	17,912,618	34.5	0.952	0.948	0.965	161.3	36.5
	GhostYoloV8n	**491**	**1,606,624**	**8.6**	**0.981**	**0.919**	**0.954**	**191.6**	**3.7**

The confusion matrix in Figure 4 offers a detailed breakdown of GhostYoloV8n's predictions. The F1-Confidence curve in Figure 5 demonstrates the balance between precision and recall at various confidence thresholds, with a high F1 score indicating that the model is highly effective in detecting both cataract and normal eye cases, underscoring its potential. The Precision-Confidence curve in Figure 6 shows consistently high precision, suggesting that when the model predicts cataracts, it is highly likely to be correct, minimizing FP. The Precision-Recall curve in Figure 7, crucial for evaluating performance on imbalanced datasets, reveals a large area under the curve, indicating that the model maintains a strong balance between precision (accurate positive predictions) and recall (identifying true positives), ensuring proper identification of cataracts without overlooking normal eyes. The Recall-Confidence curve in Figure 8 reflects the model's recall capabilities across different confidence levels, with high recall values indicating the model's reliability in detecting most cataract cases, which is vital for timely diagnosis and treatment in medical settings. Overall, these figures provide a comprehensive summary of the GhostYoloV8n model's performance across key metrics, showcasing its strong ability to accurately distinguish between normal and cataract-affected eyes. Figure 9 presents sample results demonstrating cataract or normal with the labeled box.

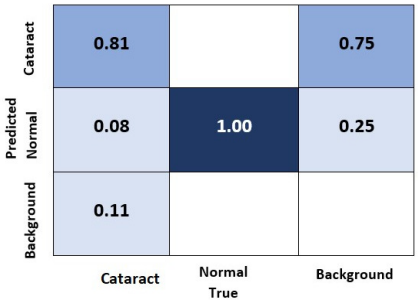

Fig. 4. Confusion matrix.

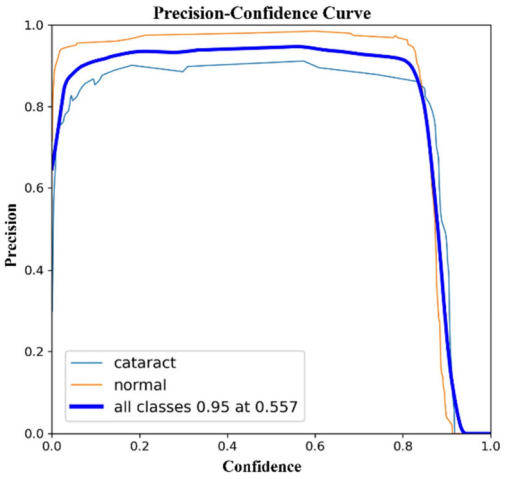

Fig. 5. F1-Confidence curve.

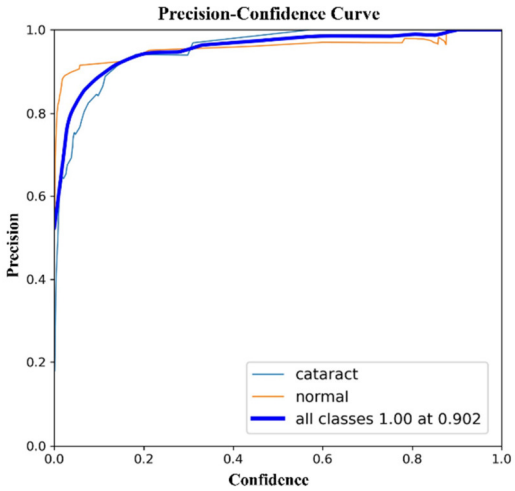

Fig. 6. Precision-Confidence curve.

Table III shows that the proposed model achieved superior performance compared to existing methods, highlighting its suitability for real-time implementation. This comparative analysis further supports the efficiency and robustness of the GhostYoloV8n model for cataract detection in practical applications. The GhostYoloV8n model leverages advanced techniques and enhanced design, improving detection capabilities and establishing itself as a reliable tool for healthcare professionals in real-world settings. The GhostYoloV8n model consistently outperforms other approaches in key metrics such as accuracy and speed. This makes it especially suitable for clinical environments, where fast and precise cataract detection is crucial for patient care.

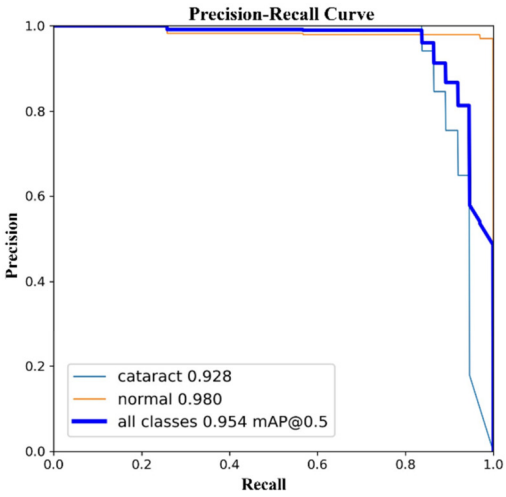

Fig. 7. Precision-Recall curve.

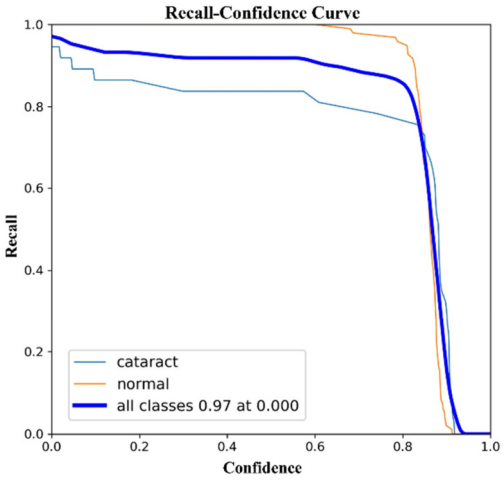

Fig. 8. Recall-Confidence curve.

Fig. 9. Results demonstrating cataract or normal within the labeled box.

TABLE III. PERFORMANCE EVALUATION OF EXISTING VS THE PROPOSED MODEL

Model	Dataset	Results
ACCV: YOLOv3 + DenseNet [18]	1,520 images from 10-second cataract videos	Accuracy: 94.00%, F1-score: 0.9388
A Lightweight Diabetic Retinopathy Detection Model (MobileNet V3 + YOLOv7) [19]	APTOS: 5,590 images	Accuracy: 98.0% and 98.4, F1-score: 93.7 and 93.1
Unified Diagnosis Framework: YOLOv3 + ShuffleNet + SVM [20]	MSLPP dataset: 16,103 anterior ocular images (4,738 pronounced cataracts, 5,346 early cataracts, 6,019 non-cataract)	Accuracy: 93.5%, F1-score: 92.3
YOLOv8: AI-Enabled Cataract Detection System [21]	1,015 images	Accuracy: 96.1%, surpassing Inception-V3, VGGNet, AlexNet, and ResNet50
Cat Eye Disease Classifier: YOLOv5 [22]	Custom dataset: 300 images (100 Cataracts, 100 Cherry Eye, 100 Glaucoma)	Accuracy: 90.00% across three diseases
Proposed Model: GhostYoloV8n	RoboFlow dataset: 1,708 images (920 Cataract, 788 Normal)	Precision: 0.981, Recall: 0.919, mAP@50: 0.954

Due to its compact architecture, featuring only 1.6 million parameters, this lightweight design not only reduces computational requirements but also speeds up processing times. GhostYoloV8n significantly enhances the efficiency of cataract detection, facilitating rapid medical intervention and making it perfect for use in resource-constrained applications. Therefore, for real-time experiments on the Jetson Nano board, GhostYoloV8n was selected to demonstrate its ability to process and display results from camera-captured images efficiently, thereby streamlining the diagnostic workflow and simplifying the detection process.

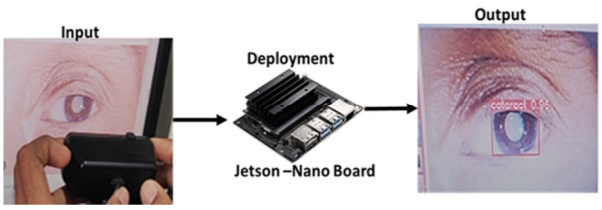

Fig. 10. Hardware implementation.

IV. CONCLUSION

GhostYOLO represents a considerable leap in cataract detection, outperforming previous models in terms of accuracy and efficiency. Using advanced object identification algorithms, this model excels at detecting early cataract signs from image data while maintaining a lightweight design with only 1.6 million parameters, making it perfect for resource-constrained situations. Unlike prior YOLO-based models, which required more computational power and struggled with real-time deployment, GhostYoloV8n incorporates GhostConv layers to improve feature extraction while reducing computational costs. This enables it to achieve the best accuracy (0.981), fastest detection speed (191.6 FPS), and lowest memory usage (3.7 MB), exceeding models such as YOLOv5, YOLOv6, and YOLOv7. A real-time implementation on the Jetson Nano board demonstrated its efficiency, with 33.5 FPS in live detection, making it appropriate for portable and cost-effective diagnostic solutions. Previous research with models such as YOLOv3 and YOLOv8 revealed accuracy levels ranging from 94% to 96.1%, but GhostYoloV8n not only increases accuracy but also dramatically decreases hardware needs, allowing smoother and faster cataract detection. This work provides a fresh and optimal technique for real-time cataract detection, addressing the limitations of current YOLO models. Future work will focus on validating GhostYoloV8n in a variety of clinical contexts and increasing its integration into telemedicine and mobile healthcare systems to improve access to early cataract diagnosis, particularly in distant and disadvantaged areas.

REFERENCES

[1] D. Lam *et al.*, "Cataract," *Nature Reviews Disease Primers*, vol. 1, no. 1, Jun. 2015, Art. no. 15014, https://doi.org/10.1038/nrdp.2015.14.

[2] L. P. L, R. Vaddi, M. O. Elish, V. Gonuguntla, and S. S. Yellampalli, "CSDNet: A Novel Deep Learning Framework for Improved Cataract State Detection," *Diagnostics*, vol. 14, no. 10, Jan. 2024, Art. no. 983, https://doi.org/10.3390/diagnostics14100983.

[3] M. V. Cicinelli, J. C. Buchan, M. Nicholson, V. Varadaraj, and R. C. Khanna, "Cataracts," *The Lancet*, vol. 401, no. 10374, pp. 377–389, Feb. 2023, https://doi.org/10.1016/S0140-6736(22)01839-6.

[4] T. Ganokratanaa, M. Ketcham, and P. Pramkeaw, "Advancements in Cataract Detection: The Systematic Development of LeNet-Convolutional Neural Network Models," *Journal of Imaging*, vol. 9, no. 10, Oct. 2023, Art. no. 197, https://doi.org/10.3390/jimaging9100197.

[5] K. Krosl *et al.*, "CatARact: Simulating Cataracts in Augmented Reality," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Recife/Porto de Galinhas, Brazil, Nov. 2020, pp. 682–693, https://doi.org/10.1109/ISMAR50242.2020.00098.

[6] D. Allen and A. Vasavada, "Cataract and surgery for cataract," *BMJ*, vol. 333, no. 7559, pp. 128–132, Jul. 2006, https://doi.org/10.1136/bmj.333.7559.128.

[7] T. Reis, V. Lansingh, J. Ramke, J. C. Silva, S. Resnikoff, and J. M. Furtado, "Cataract as a Cause of Blindness and Vision Impairment in Latin America: Progress Made and Challenges Beyond 2020," *American Journal of Ophthalmology*, vol. 225, pp. 1–10, May 2021, https://doi.org/10.1016/j.ajo.2020.12.022.

[8] I. McCormick *et al.*, "Effective cataract surgical coverage in adults aged 50 years and older: estimates from population-based surveys in 55 countries," *The Lancet Global Health*, vol. 10, no. 12, pp. e1744–e1753, Dec. 2022, https://doi.org/10.1016/S2214-109X(22)00419-3.

[9] X. Li *et al.*, "The association of cooking fuels with cataract among adults aged 50 years and older in low- and middle-income countries: Results from the WHO Study on global AGEing and adult health (SAGE)," *Science of The Total Environment*, vol. 790, Oct. 2021, Art. no. 148093, https://doi.org/10.1016/j.scitotenv.2021.148093.

[10] T. J. Ferguson and J. B. Randleman, "Cataract surgery following refractive surgery: Principles to achieve optical success and patient satisfaction," *Survey of Ophthalmology*, vol. 69, no. 1, pp. 140–159, Jan. 2024, https://doi.org/10.1016/j.survophthal.2023.08.002.

[11] H. E. Jazayeri *et al.*, "Does Early Repair of Orbital Fractures Result in Superior Patient Outcomes? A Systematic Review and Meta-Analysis," *Journal of Oral and Maxillofacial Surgery*, vol. 78, no. 4, pp. 568–577, Apr. 2020, https://doi.org/10.1016/j.joms.2019.09.025.

[12] R. Rampat *et al.*, "Artificial Intelligence in Cornea, Refractive Surgery, and Cataract: Basic Principles, Clinical Applications, and Future Directions," *The Asia-Pacific Journal of Ophthalmology*, vol. 10, no. 3, Jun. 2021, Art. no. 268, https://doi.org/10.1097/APO.0000000000000394.

[13] D. Benet and O. J. Pellicer-Valero, "Artificial intelligence: the unstoppable revolution in ophthalmology," *Survey of Ophthalmology*,

vol. 67, no. 1, pp. 252–270, Jan. 2022, https://doi.org/10.1016/j.survophthal.2021.03.003.

[14] J. P. O. Li *et al.*, "Digital technology, telemedicine and artificial intelligence in ophthalmology: A global perspective," *Progress in Retinal and Eye Research*, vol. 82, May 2021, Art. no. 100900, https://doi.org/10.1016/j.preteyeres.2020.100900.

[15] D. V. Gunasekeran and T. Y. Wong, "Artificial Intelligence in Ophthalmology in 2020: A Technology on the Cusp for Translation and Implementation," *The Asia-Pacific Journal of Ophthalmology*, vol. 9, no. 2, Apr. 2020, Art. no. 61, https://doi.org/10.1097/01.APO.0000656984.56467.2c.

[16] A. L. Pelletier, J. Thomas, and F. R. Shaw, "Vision Loss in Older Persons," *American Family Physician*, vol. 79, no. 11, pp. 963–970, Jun. 2009.

[17] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, https://doi.org/10.1007/s11042-022-13644-y.

[18] S. Hu *et al.*, "ACCV: automatic classification algorithm of cataract video based on deep learning," *BioMedical Engineering OnLine*, vol. 20, no. 1, Aug. 2021, Art. no. 78, https://doi.org/10.1186/s12938-021-00906-3.

[19] W. Sait and A. Rahaman, "A Lightweight Diabetic Retinopathy Detection Model Using a Deep-Learning Technique," *Diagnostics*, vol. 13, no. 19, Jan. 2023, Art. no. 3120, https://doi.org/10.3390/diagnostics13193120.

[20] S. Hu *et al.*, "Unified Diagnosis Framework for Automated Nuclear Cataract Grading Based on Smartphone Slit-Lamp Images," *IEEE Access*, vol. 8, pp. 174169–174178, 2020, https://doi.org/10.1109/ACCESS.2020.3025346.

[21] R. Angeline, S. Soorya, M. G. Sathyaa, and U. K. Tharun Suriya, "YOLO V8: The AI-Enabled Cataract Detection System," in *Smart Trends in Computing and Communications*, Singapore, 2024, pp. 463–474, https://doi.org/10.1007/978-981-97-1320-2_37.

[22] T. D. D. Tuason, J. Darius C. Garcia, and C. O. Manlises, "Cat Eye Disease Classifier Using YOLOv5," in *2024 7th International Conference on Information and Computer Technologies (ICICT)*, Honolulu, HI, USA, Mar. 2024, pp. 199–203, https://doi.org/10.1109/ICICT62343.2024.00037.

[23] "Eye Health3 Dataset ," *Roboflow*. https://universe.roboflow.com/cdacm/eye-health3.

[24] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[25] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 6517–6525, https://doi.org/10.1109/CVPR.2017.690.

[26] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement." arXiv, Apr. 08, 2018, https://doi.org/10.48550/arXiv.1804.02767.

[27] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv.org, Apr. 22, 2020.

[28] Ultralytics, "Comprehensive Guide to Ultralytics YOLOv5." https://docs.ultralytics.com/yolov5.

[29] C. Li *et al.*, "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications." arXiv, Sep. 07, 2022, https://doi.org/10.48550/arXiv.2209.02976.

[30] S. Ennaama, H. Silkan, A. Bentajer, and A. Tahiri, "Enhanced Real-Time Object Detection using YOLOv7 and MobileNetv3," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19181–19187, Feb. 2025, https://doi.org/10.48084/etasr.8777.

[31] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, Jun. 2023, pp. 7464–7475, https://doi.org/10.1109/CVPR52729.2023.00721.

[32] Ultralytics, "YOLOv8." https://docs.ultralytics.com/models/yolov8.