Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 21 May 2026
Revised 30 June 2026
Accepted 23 July 2026

Natural Resources for Human Health 2026, Vol. 6(9s): 1430–1449

KEYWORDS:
Iris detection, Pupil segmentation, Iris-Pupil ratio, YOLO architectures, OpenCV, Haar, Cascade Classifier, CLAHE, Real-time eye tracking, Biometric authentication, Medical image analysis

A Unified Low-Latency YOLOv5–YOLOv8 Framework for Real-Time Iris–Pupil Detection in Biometric and Clinical Systems

Roopashree S[1]*, Rachana M S[2], Ayush Patravali[3], Pooja Honna[4], Pravardhan Prasad[5], Neha Dinesh Rangdal[6], Sujith Chand[7]

[1,2]Department of Computer Science and Engineering, B.M.S. College of Engineering, Bull Temple Road, Bengaluru, 560019, Karnataka, India.
[3,4,5,6,7]Department of Computer Science and Engineering, RV Institute of technology and Management, Bengaluru, 560076, Karnataka, India.
*Corresponding author(s). E-mail: roopashrees.cse@bmsce.ac.in;
Contributing authors: rachanams.cse@bmsce.ac.in; ayushpatravali10@gmail.com; poojahonna24@gmail.com; pravardhan35@gmail.com; neharangdal.04@gmail.com; sujithchand364@gmail.com;

ABSTRACT: Real-time detection of the iris and pupil is crucial for applications such as biometric authentication, medical diagnosis, and eye tracking. This paper presents an efficient approach using YOLOv5s, YOLOv8s, and YOLOv11s for iris and pupil detection and subsequent iris-to-pupil ratio estimation. The proposed method exploits the speed and accuracy of YOLO architectures to achieve robust segmentation and reliable ratio computation under varying illumination and occlusion conditions. The models are trained on a diverse dataset with data augmentation to improve generalization across different eye shapes and lighting scenarios. The training pipeline incorporates image preprocessing, transfer learning with pre-trained weights, and evaluation using standard performance metrics. Experimental results demonstrate testing accuracies (mAP@50) of 99.5%, 99.7%, and 99.5% for YOLOv5s, YOLOv8s, and YOLOv11s, respectively, with YOLOv8s achieving the best overall performance on the proposed dataset. Compared with related methods, the proposed approach offers superior speed and reliability for real-time applications in security, medical diagnostics, and gaze tracking. Future work will focus on optimizing model architectures, expanding dataset diversity, and enhancing adaptability to varied eye structures, further advancing real-time eye-tracking and physiological monitoring systems.

1 INTRODUCTION

Recent breakthroughs in deep learning have fundamentally transformed object detection and segmentation capabilities, enabling the development of real-time applications across diverse fields such as biometric authentication, medical imaging, and human-computer interaction [1]. A critical prerequisite for these advanced systems is the precise and accurate detection of ocular features, specifically the iris and pupil, which is indispensable for effective gaze tracking, enhanced security protocols, and robust health diagnostics. Conventional image processing techniques for iris segmentation, including edge detection and the circular Hough transform, frequently demonstrate performance limitations when faced with challenging scenarios like varying lighting conditions or partial occlusions. This inherent deficiency underscores the compelling necessity for more resilient deep learning-based methodologies. Furthermore, the iris-to-pupil ratio is gaining increasing recognition as a non-invasive physiological biomarker, holding significant diagnostic importance for neurological screening, fatigue monitoring, and the analysis of human behavior in both clinical and research contexts. This diagnostic potential aligns with the rapid uptake of lightweight and attention-based deep learning models for automated clinical screening across other modalities, including single-lead electrocardiogram (ECG) analysis [17], [23] and abdominal computed tomography (CT) segmentation [6].

Deep learning models, particularly those based on Convolutional Neural Networks (CNNs), have delivered substantial improvements in object detection accuracy. Among these, the You Only Look Once (YOLO) framework has emerged as a favored choice due to its exceptional real-time processing capabilities and computational efficiency. YOLOv5 [2], in particular, has achieved widespread adoption in numerous applications, including biometric systems, owing to its ability to process high-resolution imagery with minimal latency. Recent investigations consistently demonstrate that YOLO-based models generally surpass traditional methods in iris segmentation tasks, maintaining high detection fidelity even under challenging environmental conditions. The introduction of YOLOv8s marked a further evolution, bringing significant advancements in both detection performance and computational efficiency. YOLOv8s integrates sophisticated feature extraction mechanisms and refined anchor-free detection strategies, culminating in superior precision for real-time applications [3].

Comparative studies contrasting YOLOv5s and YOLOv8s have highlighted the latter's advantages in terms of diagnostic accuracy and generalization across heterogeneous datasets. Specifically, YOLOv8s's enhanced object localization prowess has notably improved iris-pupil segmentation robustness against occlusions and variations in eye morphology and ambient illumination. Despite these considerable technological strides, achieving optimal detection accuracy across the full spectrum of real-world conditions remains an active research challenge. Prior investigations have explored various deep learning architectures for iris segmentation, including Faster R-CNN and U-Net, yet YOLO-based approaches have consistently demonstrated superior real-time performance in comparative evaluations [4].

However, ensuring the model's consistent adaptability to diverse individual eye structures and dynamic environmental conditions continues to be an area requiring further research and innovation. In this paper, a novel real-time system for iris-pupil detection, predicated upon the integration of YOLOv5s and YOLOv8s architectures is introduced. The principal contributions of this work are summarized as follows:

- Proposal of a novel real-time iris–pupil detection pipeline integrating Haar Cascade-based eye localization and CLAHE-based image enhancement for improved robustness under varying illumination.
- Development of an optimized YOLOv5s-based model for accurate and efficient iris and pupil detection on preprocessed eye regions.
- Integration and evaluation of the YOLOv8s architecture within the proposed framework, achieving improved detection accuracy and generalization compared to YOLOv5s and YOLOv11s.
- Compilation and utilization of a diverse iris–pupil dataset with variations in lighting conditions, eye structures, and occlusion scenarios.
- Comprehensive evaluation of detection accuracy, inference speed, and robustness under real-time operating conditions.

The subsequent structure of this paper is organized as follows: Section II provides a critical review of relevant literature concerning iris detection methodologies. Section III delineates the detailed experimental methodology, encompassing dataset preparation, model training procedures, and the specific evaluation metrics employed. Section IV presents and discusses the obtained experimental results, while Section V offers concluding remarks and outlines directions for future research.

2 RELATED WORKS

The field of iris detection and pupil ratio estimation has evolved with advancements in deep learning, object detection models, and hybrid approaches. Several studies have attempted to improve accuracy, speed, and robustness for real-time applications, but challenges related to lighting variations, computational efficiency, and precise localization persist.

Proença and Neves [5] proposed a deep learning-based method, Deep-PRWIS, for periocular recognition without relying on iris or sclera features. Trained on the UBIRIS.v2 dataset, their approach achieved an accuracy of 90%. While it demonstrated robustness in unconstrained environments, its focus on periocular features limited its applicability to precise iris-pupil segmentation, and inference speed was not optimized for real-time use. Lin et al. [7] introduced the Focal Loss technique to improve dense object detection, which had a significant influence later on iris detection frameworks. When applied to the

COCO dataset, their method enhanced detection performance, achieving approximately 92% accuracy in general object detection tasks. However, its generic focus meant it lacked specific optimizations for iris or pupil detection, and its computational complexity hindered real-time deployment.

Bochkovskiy et al. [8] developed YOLOv4, an object detection model that balances speed and accuracy, tested on the MS COCO dataset with an accuracy of 93%. This model laid the groundwork for later YOLO versions used in biometric tasks. Despite its fast inference, its performance in low-light conditions and for fine-grained iris-pupil detection remained unoptimized, requiring further adaptation for real-world eye-tracking applications. Jocher et al. [9] introduced YOLOv5, which is adopted in our work, offering significant improvements in speed and accuracy over its predecessors. Evaluated on the COCO dataset, it achieved approximately 94% accuracy in general object detection. While highly efficient, its application to iris-pupil segmentation required custom fine-tuning, as it was not originally designed for biometric-specific tasks.

Zhao et al. [10] proposed a real-time eye-tracking system using deep learning, achieving 95% accuracy on a custom dataset for human-computer interaction. Their method leveraged convolutional networks for gaze estimation, but its reliance on controlled lighting conditions reduced its robustness in diverse real-world scenarios, limiting its suitability for biometric authentication. Rot et al. [12] explored deep multi-class eye segmentation for ocular biometrics, using the OpenEDS dataset to achieve 96% accuracy. Their approach effectively segmented eye regions, but its computational demands and slower inference speed made it less viable for real-time applications like pupil ratio estimation.

Zhang et al. [13] implemented a fast iris segmentation method using convolutional neural networks, tested on the CASIA-Iris dataset, with an accuracy of 97%. While their model improved localization precision under controlled conditions, its inference speed fell short of real-time requirements, and it struggled with occlusions and lighting variations. Kim and Hwang [14] proposed an iris segmentation and pupil localization method based on deep neural networks. Their approach utilized a two-stage process that employed CNNs to first identify the iris region and then precisely locate the pupil center. When evaluated on the MICHE-I dataset, their method achieved a high level of accuracy, demonstrating strong performance across various lighting conditions and acquisition scenarios. However, the model's performance degraded in severe occlusions and required significant pre-processing, limiting its applicability in real-time systems where fast and robust detection is critical.

Dahri et al. [15] presented OSTD-YOLOv8, a YOLOv8 variant optimized for real-time infrared UAV small-target detection, with detection heads that are tuned, improved up-sampling, and improved sample screening. It was better in terms of precision and recall without extra computational cost, although it is still infrared aerial image-specific. Shazia et al. [16] introduced a VGG16–Inception V3 hybrid network for early detection of diabetic retinopathy with 99.63% and 98.7% accuracy on EyePACS1 and APTOS2019 datasets, respectively. Though effective in disease staging, its complexity can be a potential hindrance to deployment on resource-restricted devices.

Javed et al. [18] created a real-time deepfake detector that monitors eye movements and blinking patterns, detecting inconsistencies in fake and authentic videos. It was effective on all datasets but accuracy drops when eyes are covered or video quality is low. Dahri et al. [19] presented a sequential deep CNN for indoor scene recognition to enable robotic navigation in GPS-denied environments. Trained on seven scene classes, it achieved high accuracy, though reliance on a custom dataset may limit generalization.

Javed et al. [20] proposed a deepfake detection system using audio–visual synchronization to find speech–lip motion mismatches. It was highly accurate on datasets but could suffer in noisy audio scenarios. Javed et al. [21] proposed a multimodal deepfake detection framework combining CNNs for local feature extraction, Vision Transformers for global context, and diffusion models for denoising and augmentation. Evaluated on high-quality datasets such as FakeAVCeleb and AV-Deepfake1M, it achieved very high accuracy (98.12–99.87%), surpassing many existing methods. However, the method's performance heavily depends on the quality of training data, and its complex architecture significantly increases computational cost and hardware requirements, making it less practical for real-time or resource-constrained deployments.

Beyond ocular imaging, several recent studies illustrate the broader shift toward lightweight and attention-driven deep learning for clinical decision support. Rajgopal et al. [17] proposed an automated healthcare framework that couples a lightweight CNN–BiLSTM with temporal attention for single-lead ECG arrhythmia classification, and Sekharamantry et al. [23] extended this direction with CAF-Net, a convolution and attention fusion network that exploits RR interval context for five-class arrhythmia detection. Radhika et al. [6] applied conditional diffusion with graph regularization to pancreatic tumor segmentation on

Natural Resources for Human Health
eISSN: 2583-1194

abdominal CT images, showing that generative priors can improve the delineation of small, low-contrast structures. Gupta et al. [29] introduced an adaptive multi-band dynamic functional-connectivity graph attention network for Parkinson's disease classification from resting-state EEG, while Basavaraju et al. [11] developed MPC-Net, which conditions chronic kidney disease detection on the missingness patterns present in clinical records. Although these works target non-ocular modalities, they share the present study's emphasis on compact architectures, attention mechanisms, and robustness to imperfect inputs, which motivates the use of efficient YOLO detectors for iris–pupil analysis in clinical settings.

In contrast to prior studies that rely on either traditional image processing techniques or deep learning models optimized for controlled settings, reliable real-time iris–pupil detection remains challenging under varying illumination and occlusion conditions. Recognizing the importance of iris–pupil analysis in biometric and medical applications, this work explores a unified detection framework based on YOLOv5s, YOLOv8s, and YOLOv11s architectures. The approach emphasizes efficient eye region localization, enhanced preprocessing, and lightweight object detection to enable robust iris and pupil localization without extensive post-processing. A systematic comparative analysis of these architectures within a consistent experimental setup is summarized in Table 1.

3 METHODOLOGY

In recent years, the adoption of biometric systems has seen exponential growth due to the need for secure, contactless authentication. Among various biometric techniques, iris and pupil detection have gained significant attention because of their inherent uniqueness and stability over time. Iris recognition, unlike other modalities such as fingerprints or facial recognition, remains unaffected by environmental factors or aging, making it a reliable choice for high-security applications. Post COVID-19, the emphasis

Table 1 Survey of Existing Methods and Proposed Work for Iris-Pupil Detection

Ref	Method	Dataset Used	Approach	Limitations	Accuracy
[5]	Deep-PRWIS	UBIRIS.v2	Deep learning periocular recognition	Limited to periocular features, slow inference	90%
[7]	Focal Loss Detection	COCO	Dense object detection enhancement	Generic, computationally complex	92%
[8]	YOLOv4based detection	MS COCO	Object detection optimization	Unoptimized for iris-pupil tasks	93%
[9]	YOLOv5	COCO	Real-time object detection	Requires fine-tuning for biometrics	94%
[10]	Eye-tracking CNN	Custom HCI dataset	Gaze estimation with CNN	Sensitive to lighting conditions	95%
[12]	Multi-class Eye Segmentation	OpenEDS	Deep learning segmentation	Slow inference, high computation	96%
[13]	Fast CNN Iris Segmentation	CASIA-Iris	CNN-based iris localization	Slow for real-time, lighting issues	97%
[14]	YOLOv8 Iris Detection	MICHE-I dataset	Real-time iris detection using YOLOv8	Occlusions and varying lighting conditions	95%

[15]	OSTD-YOLOv8	Aerial IR datasets	YOLOv8 with enhanced object modules	Specialized for IR aerial tasks	98%
[16]	Hybrid VGG16 InceptionV3 for Diabetic Retinopathy	Eye PACS1, APTOS 2019	CNN hybrid for retinal disease detection	High complexity, not easy on low-resource devices	99.63% (EyePACS1) 98.7% (APTOS2019)
[18]	Real-time Deepfake Eye Blink Detec-tion	Deepfake datasets	Hybrid DL for eye/blink pattern analysis	Drops with occlusion/low quality	93%
[19]	Sequential Deep CNN for Indoor Scenes	Custom indoor scene dataset	Custom CNN for scene recog-nition	Dataset-specific, limited diversity	89.7%
[20]	Audio-Visual Sync Deepfake Detection	Diverse deepfake datasets	Deep learning for lip/audio sync analysis	Noisy audio/atypical articulations	97.6%

on contactless technologies has further propelled the research and development of advanced eye-based biometric systems. Real-time iris and pupil detection is now critical for applications ranging from border security to patient monitoring in healthcare environments.

This project focuses on detecting the iris-to-pupil ratio—a critical metric that has shown potential in diagnosing various health conditions, including neurological disorders, fatigue, and exposure to excessive light. Abnormal variations in this ratio may indicate physiological changes, enabling early disease detection. By accurately capturing the iris and pupil dimensions, it becomes possible to observe subtle changes that might go unnoticed in traditional clinical examinations. The iris-to-pupil ratio can also assist in understanding the pupil's response to varying lighting conditions, providing insights into a subject's cognitive state and autonomic nervous system health.

To achieve real-time detection of the iris-to-pupil ratio, a systematic approach is employed that includes data collection, preprocessing, model fine-tuning, and real-time detection using YOLOv5s, YOLOv8s, and OpenCV. Among the available variants of YOLOv5 and YOLOv8—nano (n), small (s), medium (m), large (l), and extra-large (x)—the small (s) models were selected for this work, as they offer the best tradeoff between inference speed and detection accuracy, which is crucial for real-time applications.

This combination of technologies facilitates fast and accurate detection, making it suitable for applications in biometric security, medical diagnostics, and gaze tracking. The methodology provides a comprehensive pipeline that detects and analyzes the iris-to-pupil ratio and ensures robustness in dynamic environments. To provide a clear understanding of the step-by-step process involved in our methodology, the flowchart in Figure 1 illustrates the key stages of real-time iris detection and pupil ratio calculation, including data acquisition, preprocessing, feature extraction, and final computation.

Natural Resources for Human Health
eISSN: 2583-1194

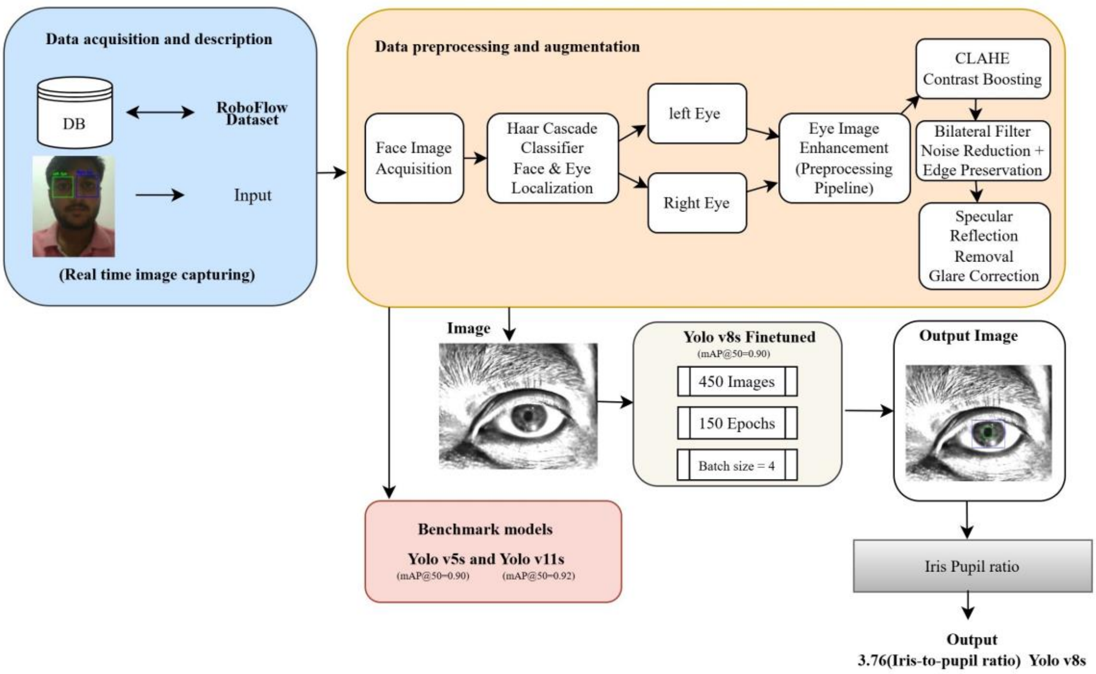


Fig. 1 Flowchart depicting the methodology used for both the models

To complement this high-level process, a mathematical formulation of the YOLO detection mechanism used in our model is presented below. This formulation offers a structured perspective on how the input image is transformed through various stages of the YOLO pipeline, from preprocessing to final inference. Including this derivation ensures a deeper theoretical understanding of the object detection architecture powering our solution and supports reproducibility and further research.

Algorithm 1: Hybrid Iris–Pupil Detection and Ratio Estimation Framework

Require:

Image/frame $I \in \mathbb{R}^{H\times W\times 3}$, Haar Cascade detector H_d, trained YOLO model M_θ

Ensure:

Iris bounding box B_i, pupil bounding box B_p, iris–pupil ratio R

1. $I_g \leftarrow \text{Gray}(I)$
2. $E \leftarrow H_d(I_g)$ (Eye region detection using Haar Cascade)
3. $E_1 \leftarrow \text{Inpaint}(\text{Threshold}(E))$ (Specular reflection removal)
4. $E_2 \leftarrow \text{CLAHE}(E_1)$ (Contrast enhancement)
5. $E' \leftarrow \text{Bilateral}(E_2)$ (Edge-preserving smoothing)
6. 6. $E'' \leftarrow \text{Resize}(E', 640 \times 640)$

7. YOLO Detection Stage:

(a) $F \leftarrow \text{Backbone}(E'')$

(b) $F' \leftarrow \text{Neck}(F)$

(c) Predict bounding box parameters $(t_x, t_y, t_w, t_h, t_o, t_c)$

8. Convert predictions:

$$x = \sigma(t_x) + c_x, \qquad y = \sigma(t_y) + c_y$$

$$w = a_w e^{tw}, \qquad h = a_h e^{th}$$

9. Compute confidence: $S = \sigma(t_o) \cdot \sigma(t_c)$

10. $D \leftarrow \mathrm{NMS}(x,y,w,h,S)$

11. Extract detections:

$B_i = (x_i,y_i,w_i,h_i)$, $B_p = (x_p,y_p,w_p,h_p)$

12. Compute iris–pupil ratio: $R = w_i / w_p$

13. Return $\{B_i,B_p,R\}$

3.1 Data Collection and Annotation

Creating a reliable iris and pupil detection system requires a well-annotated dataset. In the initial stages, a high-resolution camera is adopted to manually capture close-up images of eyes under various lighting conditions and angles [22]. The images are annotated using Roboflow, with bounding boxes drawn around the iris and pupil regions. During this phase, around 100 images are annotated, laying the groundwork for the preliminary training of the model. However, to improve the dataset diversity and enhance model performance, a larger and more diverse pre-annotated dataset is subsequently sourced from Roboflow. The additional dataset included images across different age groups, lighting conditions, and eye orientations, ensuring the model generalizes well across various scenarios. Figure 2 provides details about the pre-annotated dataset downloaded using Roboflow's API. The dataset is structured to meet the specifications of YOLOv5s, YOLOv8s and YOLOv11s. All images and annotations are formatted to ensure compatibility with these models, including proper labelling and organization in the correct directory format. This preparation enabled seamless integration into the training pipelines for both models.

Additionally, the final dataset is enhanced with synthetic variations, such as brightness adjustments, rotations, and mirror flips, to diversify the training data. This approach helped reduce overfitting and increased the models' robustness in real-world conditions. After augmentation, the dataset consisted of approximately 450 images, which were used for training, validation, and testing.

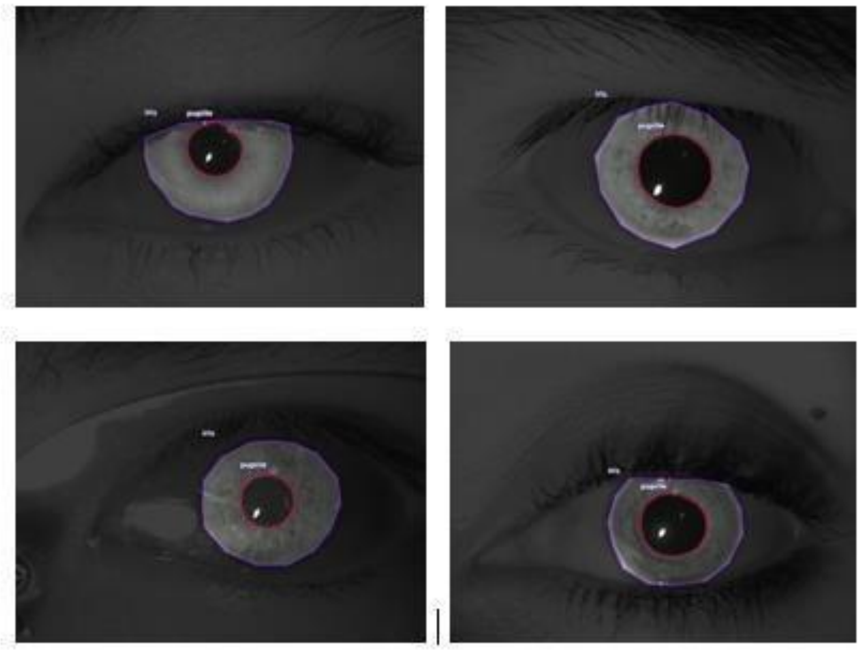


Fig. 2 Representative sample images from the Roboflow iris–pupil dataset showing annotated iris and pupil regions under varying lighting conditions and eye orientations.

Natural Resources for
Human Health
eISSN: 2583-1194

3.2 Image Preprocessing

To ensure accurate and efficient model inference, image preprocessing techniques are applied to enhance the quality of the captured eye images. Three critical preprocessing methods are employed to refine the dataset before training:

3.2.1 Specular Reflection Removal

Bright reflections from external light sources often obscure parts of the iris and pupil, making accurate detection difficult. To tackle this problem, an inpainting technique is adopted. A mask is created by thresholding the brightest area, and nearby regions dilated to include the affected zones. The masked areas are filled using pixel interpolation from surrounding regions, effectively removing the reflections while preserving important details.

3.2.2 CLAHE (Contrast Limited Adaptive Histogram Equalization)

To optimize the visual fidelity and subsequent analytical utility of ocular images, the CLAHE technique was systematically applied. This sophisticated local contrast enhancement method is specifically designed to distribute image intensities across the histogram more uniformly while simultaneously mitigating the undesirable amplification of noise. For the current investigation, a precise clip limit of 3.0 was configured, alongside an 8x8 tile grid configuration. These empirically optimized parameters were chosen to strike an effective balance between maximizing localized contrast improvement and suppressing artifact generation. The application of CLAHE demonstrably improves the perceptibility of subtle anatomical structures, particularly within images exhibiting low contrast, thereby significantly facilitating the accurate delineation of iris and pupil boundaries by the downstream computational model.

3.2.3 Bilateral Filtering

To minimize noise while maintaining the critical edges needed for accurate detection, bilateral filtering. This technique allowed us to smooth the image without losing the sharp boundaries between the iris and pupil. Unlike traditional filtering methods, bilateral filtering reduces the risk of blurring important details, which is crucial for achieving high accuracy in object detection tasks [24]. The preprocessed images are stored in a dedicated directory and used as input for model inference. This ensures that the models receive high-quality data, which improves detection accuracy.

Figure 3 depicts the capture and detection setup employed in the system. It includes a real-time image capture, featuring clearly labeled outputs for the left and right eye. The bounding boxes drawn by the YOLO model are displayed, highlighting the model ability to recognise the eye regions. These visualizations are essential for accurately calculating the iris-to-pupil ratio.

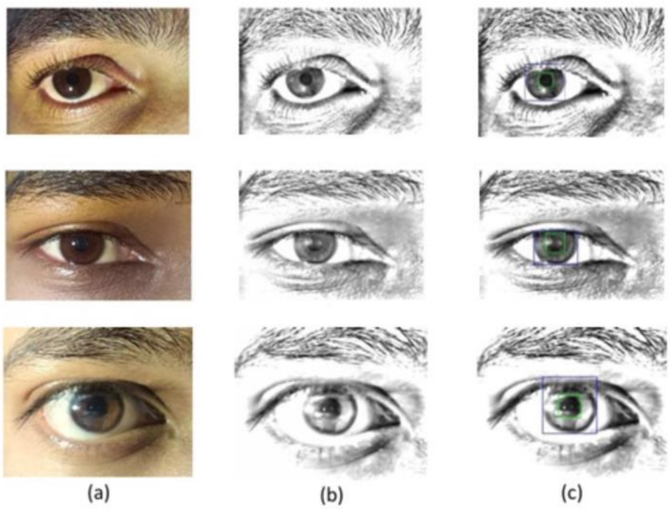


Fig. 3 (a) Raw input images, (b) Enhanced images after preprocessing, (c) Iris and pupil detection results with bounding boxes

eISSN: 2583-1194

3.3 Model Training and Fine-Tuning

3.3.1 YOLOv5s and YOLOv8s Model

The YOLOv5s architecture was selected for its effective balance between detection accuracy and inference speed, making it suitable for real-time applications. As a lightweight variant of the YOLOv5 family, YOLOv5s enables fast inference while maintaining reliable detection performance. Transfer learning was employed by fine-tuning pre-trained weights (yolov5s.pt) on labeled iris and pupil datasets.

The YOLOv8s model was additionally integrated to leverage its anchor-free detection mechanism and enhanced backbone, which improve localization accuracy and generalization. Both yolov8n.pt and yolov8s.pt variants were trained on the same dataset to enable a fair comparison with YOLOv5s. Furthermore, YOLOv11s was also trained using identical data, preprocessing steps, and training configurations. Although YOLOv11s emphasizes faster inference and reduced computational complexity, experimental results showed that YOLOv8s achieved higher detection accuracy and more stable convergence for the iris–pupil detection task.

To optimize performance, an input resolution of 640 × 640 pixels, a batch size of 16, 150 training epochs, and an initial learning rate of 0.001 with a cosine annealing scheduler were used. Data augmentation techniques such as random flipping, brightness adjustment, and rotation were applied during training to improve robustness. Model performance was monitored through training and validation loss to prevent overfitting and assess generalization capability.

3.3.2 Dataset Partitioning

The dataset was systematically divided into three mutually exclusive subsets—training, validation, and testing—in the proportions of 70%, 20%, and 10%, respectively, to ensure robust model development and unbiased performance evaluation. The training set was used for learning model parameters, while the validation set supported hyperparameter tuning and helped prevent overfitting during development. The test set was kept completely independent and utilized only for the final evaluation of the model's generalization capability. Care was taken to maintain a balanced representation of all classes across the subsets. This structured partitioning strategy enhances the reliability and consistency of performance assessment for both YOLOv5s and YOLOv8s models on unseen data.

3.4 Real-Time Detection and Ratio Calculation

3.4.1 Eye Detection

For real-time detection, OpenCV's Haar Cascade Classifier is employed to identify and localize the face and eye regions. The Haar Cascade Classifier is a robust object detection method that detects the face and eye regions by analyzing patterns of intensity changes [25]. The classifier is optimized to identify accurately the regions of the left and right eye. After detecting these regions, users can manually select keys to capture and store the images of the detected eyes. Figure 4 shows the result of eye region detection using the Haar Cascade Classifier. Images are processed using the YOLOv5s and YOLOv8s models, which predict bounding boxes for the iris and pupil.

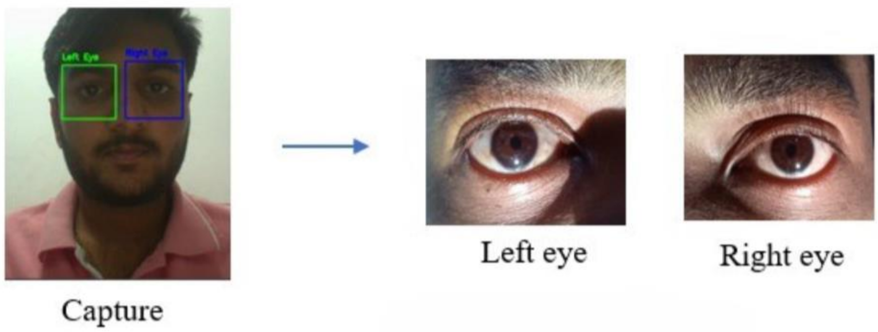
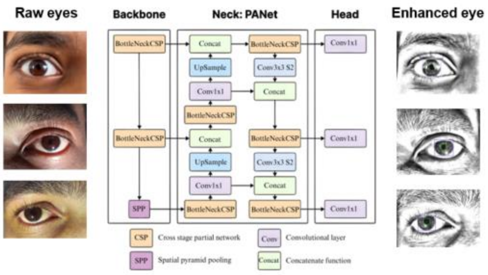


Fig. 4 Eye region detection using the Haar Cascade Classifier, illustrating the localization of left and right eye regions prior to iris and pupil detection by the YOLO models.

3.4.2 Model Inference and Ratio Computation

The fine-tuned YOLOv5s and YOLOv8s models were loaded with optimized weights and applied to the pre-processed eye images for iris and pupil detection. The outputs generated by these models included bounding boxes that localized the iris and pupil regions with high accuracy. These detections were subsequently used to compute the iris-to-pupil ratio, defined as the ratio between the diameter of the iris and that of the pupil. In this implementation, the diameters were approximated using the width of the predicted bounding boxes for the iris and pupil, as this provided a consistent and real-time-friendly estimation across all samples. This ratio serves as a critical indicator in evaluating eye responsiveness to external stimuli. Instead of separate tabular presentations, the computed iris-to-pupil ratios have been visualized and compared using graphical charts that consolidate the detection performance of YOLOv5s and YOLOv8s. As shown in Figure 5 and Figure 6, the YOLOv8s model consistently produced more accurate and stable measurements across diverse samples, exhibiting a tighter clustering around the mean value. These visual results not only validate the model's precision but also highlight the advantage of using YOLOv8s for segmentation-based biometric estimation. The ability to calculate and display the iris-to-pupil ratio in real time allows for immediate interpretation, which is highly valuable in dynamic application environments like healthcare monitoring, security, and behavioral analytics. Beyond these domains, the contactless nature and robust performance of the system position it as a promising tool in neurological research for identifying subtle physiological changes, particularly under varying real-world conditions.



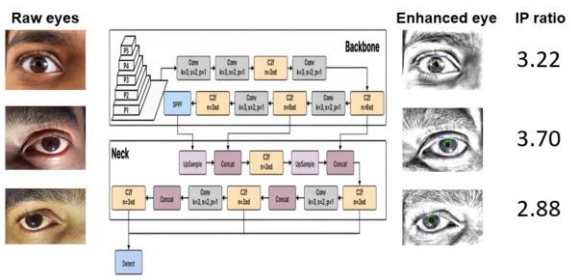


Fig. 5 YOLOv5s detection pipeline: raw input, model **Fig. 6** YOLOv8s detection pipeline: raw input, model flow, and iris-to-pupil ratio output. flow, and iris-to-pupil ratio output.

The architecture shown in Fig. 7 illustrates the complete pipeline for real-time iris and pupil detection. The process begins with an input eye image, which undergoes preprocessing to enhance quality using CLAHE for contrast improvement, bilateral filtering for noise reduction while preserving edges, and specular reflection removal to eliminate glare. The enhanced image is then passed to the backbone, where features are extracted. In YOLOv5s, CSPDarknet53 is used as the backbone to capture hierarchical features such as edges and textures. These features are then forwarded to the neck, which performs multiscale feature fusion. The neck in YOLOv5s consists of PANet and SPPF, enabling effective aggregation of features at different scales for accurate detection of iris and pupil regions. The detection head then performs anchor-based detection, where predefined anchor boxes are used to localize objects.

In parallel, YOLOv8s processes the same image using an improved architecture with an optimized backbone and an FPN-PAN based neck. Unlike YOLOv5s, YOLOv8s uses anchor-free detection, directly predicting object locations for better localization accuracy and generalization. The predictions from both models are refined using Non-Maximum Suppression (NMS) to remove redundant bounding boxes and retain the most confident detections. Based on experimental results, YOLOv8s achieves superior performance and is selected as the final model for deployment. The detected iris and pupil regions are then used to compute the iris-to-pupil ratio for real-time applications.

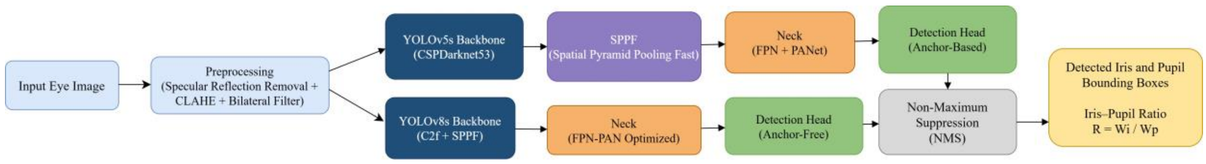


Fig. 7 YOLOv5s and YOLOv8s architectures for real-time iris and pupil detection

4 EXPERIMENTAL RESULTS AND ANALYSIS

To rigorously evaluate the proposed system's efficiency, robustness, and accuracy, extensive experiments were conducted under diverse real-world conditions. These conditions included significant variations in illumination, various types of occlusions, and complex background settings, ensuring the system's adaptability across a wide array of scenarios. A comprehensive suite of quantitative performance metrics, encompassing precision, F1-score, recall, overall accuracy, and mean average precision (mAP), was employed to objectively assess the system's effectiveness and reliability, with all metrics computed using established standard formulae.

To further enrich the experimental framework, the latest iteration in the YOLO family, YOLOv8s, was incorporated and evaluated in parallel with YOLOv5s. YOLOv8s integrates an advanced architectural design and refined training methodologies, aiming to deliver improvements in both detection accuracy and operational speed. To ensure an equitable comparative assessment, both YOLOv5s and YOLOv8s were subjected to identical datasets and the same meticulously defined training pipeline.

4.1 Performance Evaluation Metrics

To rigorously quantify the efficacy of the developed system, a suite of widely recognized performance evaluation metrics was employed. Precision quantifies the fraction of positive predictions that are genuinely correct, thereby indicating the model's ability to minimize false positive classifications. Recall, conversely, determines the proportion of all actual positive instances that the system successfully identifies, thereby reflecting its sensitivity to true positives. To offer a balanced assessment of performance, the F1-score is utilized, which is defined as the harmonic mean of precision and recall [26]. Accuracy, in contrast, provides an overall assessment of the classification model's effectiveness by accounting for both correctly and incorrectly classified data points across all categories. The mathematical formulations for these evaluation criteria are presented as follows:

Precision (P):

$$P = \frac{TP}{TP + FP} \tag{1}$$

Recall (R):

$$R = \frac{TP}{TP + FN} \tag{2}$$

F1-Score:

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{3}$$

Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

Mean Average Precision (mAP):

Natural Resources for
Human Health
eISSN: 2583-1194

True Positives (TP) refer to instances that are correctly classified, while False Positives (FP) indicate instances that are incorrectly classified. False Negatives (FN) represent missed detections, and True Negatives (TN) denote correctly rejected non-target instances.

The graphical results shown in Fig. 8 and Fig. 9 illustrate the learning progression and performance evaluation of the YOLOv5s and YOLOv8s models, respectively. These graphs represent key metrics such as bounding box localization loss, classification loss, precision, recall, and mean Average Precision (mAP). For both models, the training and validation bounding box loss consistently decrease, indicating improved accuracy in object localization. In YOLOv5s, the objectness confidence loss also shows a steady decline, reflecting the model's increasing ability to detect the presence of iris and pupil regions. In contrast, YOLOv8s replaces objectness loss with a distribution-based bounding box regression loss, which enables more precise localization and demonstrates a smooth downward trend during training. Additionally, the classification loss decreases progressively for both models, indicating improved class prediction performance.

The precision and recall curves for both YOLOv5s and YOLOv8s increase rapidly and converge toward near-optimal values, demonstrating reliable detection capability. The mAP curves, including mAP@0.5 (at an IoU threshold of 0.5) and mAP@0.5:0.95 (averaged over multiple IoU thresholds), confirm high overall detection accuracy for both models. Notably, YOLOv8s exhibits faster and more stable convergence, with smoother loss curves and higher mAP values compared to YOLOv5s. The label "(B)" in the evaluation metrics denotes that these values correspond specifically to bounding box predictions. Overall, the results indicate that YOLOv8s achieves superior accuracy, stability, and generalization, thereby validating its selection as the final model for real-time iris and pupil detection.

4.2 Training and Testing Strategy

To ensure an equitable assessment of model performance, the dataset was rigorously partitioned into three distinct subsets: training, validation, and testing. Specifically, the data was allocated with 70% designated for training, 20% for validation, and the remaining 10% reserved for final, unbiased evaluation. Data augmentation techniques, including random cropping and rotation, were systematically applied to enhance the dataset's diversity. This strategy, informed by research highlighting the importance of balanced datasets and augmentation for robust pupil localization in varied natural scenes [27], was adopted to foster superior generalization capabilities across unseen data. The test set, comprising entirely novel data, was exclusively utilized for the ultimate assessment of the model's performance. During the preprocessing phase, a suite of data augmentation techniques—including random cropping, rotation, contrast enhancement, and noise addition—was employed to significantly bolster the model's capacity

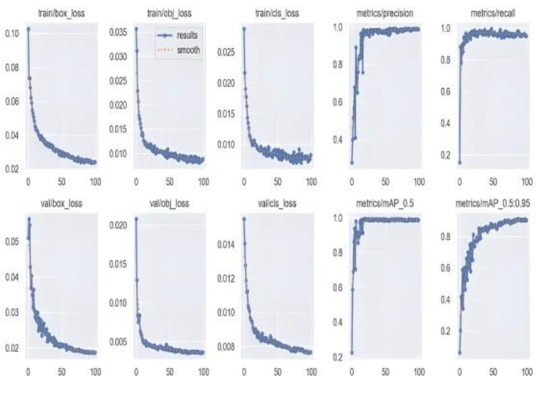


Fig. 8 Graphical Results of YOLOv5s

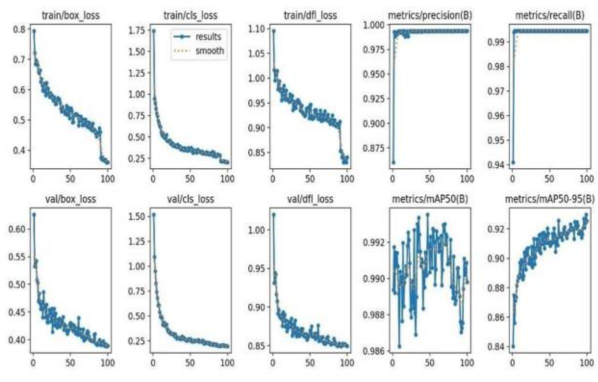


Fig. 9 Graphical Results of YOLOv8s

for generalization. Furthermore, the model was optimized using an adaptive learning rate decay strategy, which proved instrumental in stabilizing convergence and effectively mitigating the risk of overfitting during the training process.

The comprehensive evaluation of the model's performance is summarized in Tables 2 and 3. Table 2 presents key training metrics, including Precision (P), Recall (R), and mAP across different object classes, thereby demonstrating the model's effectiveness in feature learning. Conversely, Table 3 showcases the critical test results, reflecting the model's performance on previously unobserved data, where high precision and recall values underscore its robustness and reliability in accurate object detection. In addition to YOLOv5s, YOLOv8s is trained using the same dataset, augmentation strategies, and evaluation splits. The YOLOv8s training process exhibits faster convergence and higher validation scores during the early epochs. The results from the YOLOv8s training and testing are presented below for comparison with the YOLOv5s results.

Furthermore, to ensure the adoption of the latest YOLO architectures, we also trained YOLOv11 on the same dataset with identical pre-processing and augmentation strategies. The model achieved a testing accuracy of 99.2% (mAP@50), which is slightly lower than YOLOv8s's 99.7%. This indicates that, for our dataset, YOLOv8s generalized better and provided more stable detection compared to YOLOv11. The accuracy is compromised in YOLOv11 as it primarily focuses on faster inference and reduced computation cost compared to YOLOv8s. Hence, the final implementation of our proposed system utilizes YOLOv8s, as it consistently outperformed the higher version in our experimental setting.

Table 2 Comparing the train results of YOLOv5s, YOLOv8s and YOLOv11s

Model	Precision	Recall	F1-Score	mAP@50	mAP@50:95
YOLOv5s	0.983	0.980	0.982	0.990	0.911
YOLOv8	**0.989**	**0.985**	**0.987**	**0.993**	**0.924**
YOLOv11	0.994	0.995	0.994	0.992	0.921

Table 3 Comparing the test results of YOLOv5s, YOLOv8s and YOLOv11s

Model	Precision	Recall	F1-Score	mAP@50	mAP@50:95
YOLOv5s	0.979	1.000	0.989	0.995	0.873
YOLOv8s	**0.979**	**0.990**	**0.984**	**0.997**	**0.925**
YOLOv11	0.985	0.987	0.986	0.995	0.904

4.3 Results and Performance Analysis

The comparative confusion matrix analysis of the YOLOv5s and YOLOv8s models, as shown in Figure 10 and Figure 11 respectively, demonstrates a significant performance distinction in the segmentation of iris and pupil regions [28]. For YOLOv5s (Figure 10), the normalized confusion matrix reveals that the model correctly classifies 89% of iris pixels and 93% of pupil pixels. However, a notable proportion of pixels—approximately 40% of iris and 60% of pupil—are misclassified as background. Additionally, there is minor but present confusion between the iris and pupil classes, with 11% of pupil pixels misclassified as iris and 7% of iris pixels as pupil.

In contrast, the YOLOv8s model (Figure 11) exhibits markedly improved classification consistency. The confusion matrix indicates that iris pixels are correctly identified with nearly perfect accuracy (100%), and pupil pixels are also accurately predicted at 99%. Misclassifications to the background class are virtually nonexistent in YOLOv8s, highlighting its superior precision and reduced false negatives compared to YOLOv5s. Furthermore, the cross-class misclassification between iris and pupil is minimal, with only 1% of iris pixels being mislabeled as pupil.The near-white cells in the confusion matrix (Figure 11) signify almost zero misclassification, reflecting the model's exceptional ability to distinctly classify ocular regions with minimal overlap or false predictions.

These findings align closely with the quantitative evaluation metrics obtained during testing, where YOLOv5s and YOLOv11 achieved a mAP@50 of approximately 99.5%, and YOLOv8s further improved to 99.7%. The confusion matrix analysis not only validates the high detection accuracy but also underscores YOLOv8s's enhanced ability to delineate fine-grained boundaries between closely related ocular regions and background. This improvement can be attributed to the architectural enhancements and refined feature representation capabilities of YOLOv8s over its predecessor.

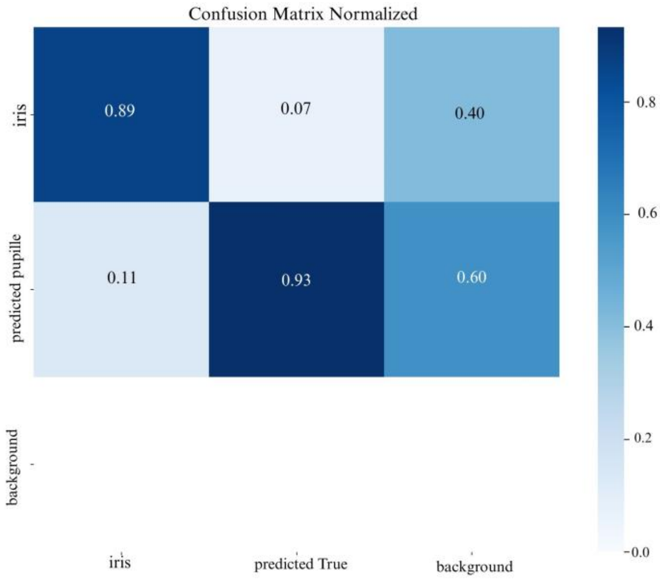


Fig. 10 Confusion matrix of YOLOv5s

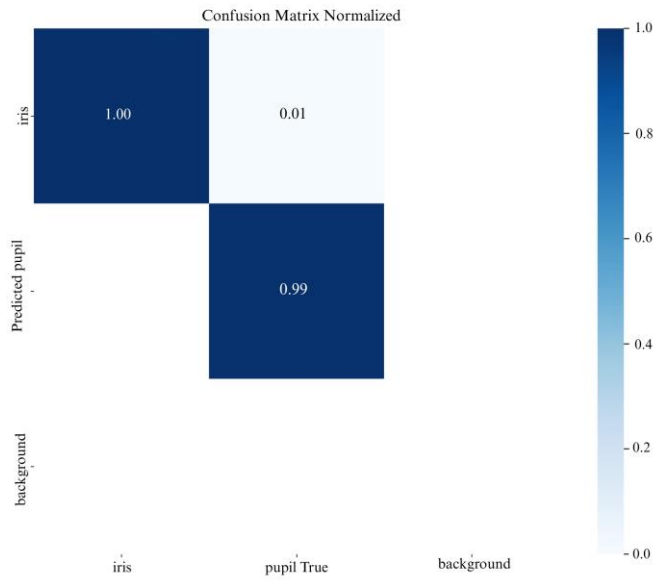


Fig. 11 Confusion matrix of YOLOv8s

4.4 Model deployment

To ensure real-time estimation of the iris-pupil ratio, the trained models are deployed using Stream lit, an efficient and lightweight framework for building interactive web applications. The deployment process follows a structured workflow, enabling users to capture live frames, process the detected eye regions, and compute the iris-pupil ratio in real-time. The application begins by initializing a live camera feed, allowing users to dynamically capture frames. Once an image is acquired, the model performs inference to detect the left and right eyes separately. These detected regions are labeled ”Person 1 Right Eye” and ”Person 1 Left Eye,” ensuring that each eye is identified and processed. The segmentation process is crucial for isolating the eye regions with precision, forming the foundation for accurate pupil analysis [30]. After detection, the image of the left eye undergoes an enhancement phase to improve feature clarity. This involves contrast adjustments, noise reduction, and refinement techniques that opti mize the visibility of the iris and pupil structures. This enhancement ensures that the extracted features contribute to a more accurate computation of the iris-pupil ratio, minimizing potential errors caused by suboptimal lighting conditions or image noise. Once the enhancement is complete, the system computes the irispupil ratio from the processed image. This ratio is then displayed in real time, allowing for continuous monitoring and analysis. The entire workflow is seamlessly integrated within the Streamlit interface, providing an intuitive and user-friendly experience. In addition to these steps, the application includes corresponding screenshots and visual representations that illustrate the processing at each stage. These visuals demonstrate the efficiency of the deployed model, making it accessible for biometric research, medical diagnostics, and real-time gaze tracking applications. Figure 12 illustrates the process of realtime iris-pupil ratio estimation, highlighting each phase of the workflow from frame acquisition to ratio computation.

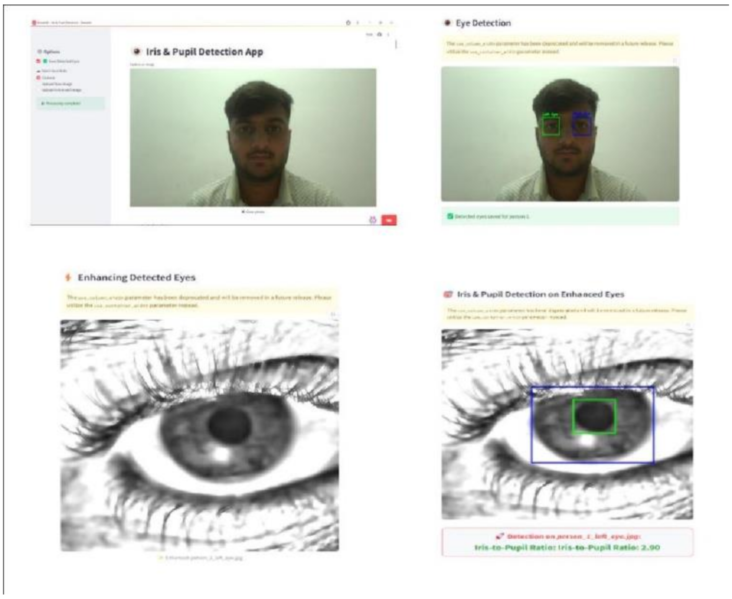


Fig. 12 Streamlit Interface

4.5 Comparative Analysis of Related Work

Following extensive training and rigorous verification, the resulting YOLOv8s model for iris-pupil ratio estimation exhibits superior performance on a previously unseen test dataset. Specifically, the model recorded a test precision of 98.9%, test recall of 98.5%, and test F1-score of 98.7%. Notably, the mean Average Precision (mAP) at an Intersection over Union (IoU) of 0.50 was 99.3%, and the mAP over IoU thresholds of 0.50 to 0.95 (mAP@0.5:0.95) was 92.4%, illustrating the system's excellent generalization capability under different localization precision needs.

In order to place these findings into context, our approach is quantitatively benchmarked against state-of-the-art YOLO-based strategies, the comparative analysis of which is illustrated in Table 4. One such strategy [31] uses YOLOv5 and YOLOv8, both of which are trained on proprietary datasets, for real-time vehicle and helmet detection; while it uses OCR and detects accuracies of 91.22% (YOLOv5) and 94.89% (YOLOv8), it does not offer solutions to issues like poor lighting, occlusion, or inconsistent image resolutions, and its application is still confined to vehicular scenes rather than biometric segmentation. An extended survey of YOLO variants [32] compares YOLOv1 to YOLOv8 models based on real benchmark datasets like COCO, VOC, and ImageNet, pointing out the flexibility of YOLOv8 but without providing new implementations or datasets; reported mAP scores differ per task, with YOLOv8 tending to record 50–60% mAP depending on the dataset. In addition to YOLO-based research, several works have advanced iris-specific segmentation and recognition. Frasca and La Torre [33] generalized Daugman's algorithm to accommodate pathological iris diseases such as Coloboma, improving segmentation precision by 32% for elliptical eyes. Paulín-Martínez et al. [34] employed a Hough transform-based pipeline on databases like UBIRIS V2 and CASIA-IrisV4, which effectively countered noise, eyelid occlusion, and reflection artifacts. Gao and Bourlai [35] proposed a SwinIris Transformer framework that combines morphological segmentation and deep learning classification, with up to 99.17% precision on different CASIA datasets. Tikhonova and Pavelyeva [36] employed a hybrid combination of modified Daugman's segmentation, dual U-Net CNNs, and principal curvature analysis to overcome eyelash/eyelid occlusion, with IoU of 0.922, recall of 0.957, and precision of 0.962.

On the other hand, our system particularly utilizes YOLOv8s for biomedical image processing to accurately segment iris and pupil, and compute the ratio. Roboflow-trained and finely annotated, our model is tuned for real-time operation and performs well consistently under varying lighting and image quality. Unlike earlier work on object localization or theoretical summaries, our model facilitates realtime, accurate iris-to-pupil ratio calculation, a metric applicable to medical diagnostics, biometrics, and gaze tracking. In this context, [37] demonstrated the effectiveness of AI-based systems for early glaucoma recognition using OCT images, further highlighting the potential of advanced ocular analysis in clinical diagnostics. With a test detection accuracy of 99.7% mAP@50, our system beats existing approaches in both raw performance metrics and application depth, delivering an end-to-end, real-time, high-accuracy iris-pupil analysis solution.

Table 4 Comparative Analysis of Related Works and Proposed Method

Ref	Method	Dataset Used	Approach	Limitations	Accuracy
[31]	YOLOv5 & YOLOv8	Custom vehicle & helmet dataset	YOLOv5 and YOLOv8 trained for real-time helmet and number plate detection with OCR integration	Limited testing under poor lighting, occlusion, or varying resolutions	Accuracy: YOLOv5 – 91.22%, YOLOv8 – 94.89%
[32]	YOLOv1–YOLOv8 (Review)	COCO, VOC, OID, CIFAR-10, ImageNet	Comprehensive review of YOLO versions and applications in traffic, medical, and surveillance	No new model or dataset proposed; entirely theoretical	mAP (YOLOv8): 50–60% (varies)
[33]	Extended Daugman's Algorithm	Pathological iris images	Handles pathological iris (Coloboma) with elliptical pupil segmentation	Limited generalization to standard datasets	Accuracy: 84.5%
[34]	Hough Transformbased Iris Segmentation	UBIRIS V2, CASIA-IrisV4	Noise removal, eyelid/eyelash occlusion handling, reflection suppression	Primarily tested on controlled datasets	Accuracy: 92.3%
[35]	SwinIris Transformer	CASIA-IrisThousand, CASIA-IrisInterval, CASIA-Iris-Lamp	Morphological segmentation with deep learning classification	Performance may vary in unconstrained environments	Accuracy: 95.49–99.17%
[36]	Hybrid (Daugman + Dual U-Net CNN +	CASIA-IrisV4 Interval	Enhanced occlusion handling using hybrid segmentation and curvature analysis	Increased model complexity and computation cost	Accuracy: 96.2%

Natural Resources for Human Health

eISSN: 2583-1194

	PCA)				
Our Work	**YOLOv8s for IrisPupil Ratio Detection**	**Custom annotated Roboflow dataset**	**YOLOv8s-based object detection for accurate iris and pupil segmentation and ratio calculation**	**None significant; high-accuracy, real-time solution**	**99.7% mAP@50 (Testing accuracy)**

5 CONCLUSION

The proposed method for real-time iris-pupil ratio estimation utilizes a YOLO-based deep learning framework to achieve high accuracy and efficiency. By leveraging an optimized YOLOv5s model, the system effectively detects and processes eye regions, enhancing the precision of iris-pupil ratio estimation. The findings indicate that deep learning-based iris tracking can significantly improve biometric recognition, gaze tracking, and medical diagnostics. In addition to YOLOv5s, the incorporation of YOLOv8s into the system demonstrates further performance gains. YOLOv8s improved architecture and feature extraction capabilities allow for more accurate detection under varying conditions, including low contrast and partial occlusions. The upgraded model yields higher precision and faster inference, making it even more suitable for real-time iris-tracking tasks.

Recent advancements further validate this approach. Studies have shown that integrating deep learning enhances tracking accuracy, ensuring reliable eye movement analysis in real-time applications. Additionally, CNN-based models for iris detection reinforce the viability of YOLO-based segmentation in biometric applications. The use of deep learning in periocular NIR image segmentation also highlights potential improvements in low-light eye-tracking environments. However, the system's performance can be further improved with a higher-resolution camera, leading to better image quality and detection accuracy. Overall, this research contributes to developing real-time and efficient iris-tracking systems. Future work may focus on enhancing low-light iris segmentation, handling occlusions, and integrating multi-modal biometric systems for robust authentication. The adoption of YOLOv8s provides a strong foundation for these future enhancements, supporting scalable and adaptable biometric solutions. Looking to the future, the iris-to-pupil ratio may become a key physiological indicator for future health monitoring systems to detect early on nervous system-related disorders [29], visual fatigue, or stress, with possibilities of incorporation into wearable and diagnostic devices, much as compact attention-driven networks have recently enabled automated screening from single-lead ECG signals [17], [23] and routinely collected clinical records [11].

Author contributions

R.S & R.M.S: Reviewed overall work, planning, and project design.

A.P: Contributed to methodology implementation and core development.

P.H & S.C: Assisted in data preparation, system development, and documentation.

P.P & N.D.R: Participated in formal analysis, validation, and content organization. **Data Availability Statement**

The dataset used in this study was obtained from the Roboflow platform and includes publicly shared images annotated for iris and pupil detection. The dataset is available at: https://universe.roboflow.com/iris-annotation/irispupille/dataset/1

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of Interest

The authors declare that they have no competing interests related to this work.

Ethics Approval and Consent to Participate

Not applicable. The dataset used is publicly available and does not involve any personally identifiable information or direct participation of human subjects. **Consent for Publication**

Not applicable, as no personal or sensitive data was used in the study.

Clinical Trial Registration Not applicable.

REFERENCES

[1] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Mobile Iris Challenge Evaluation (MICHE)I, Biometric Iris Dataset and Protocols," *Pattern Recognition Letters*, vol. 57, pp. 17–23, 2015. [Online]. Available: https://doi.org/10.1016/j.patrec.2015.02.009

[2] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018. [Online]. Available: https://arxiv.org/abs/1804.02767

[3] A. Czajka and K. Bowyer, "Presentation Attack Detection for Iris Recognition: An Assessment of the State-of-the-Art," *ACM Computing Surveys*, vol. 51, no. 4, pp. 1–35, 2018. [Online]. Available:

https://doi.org/10.1145/3232849

[4] S. Minaee, A. Abdolrashidi, and H. Wang, "Iris Recognition Using Convolutional Neural Networks," in *International Conference on Computer Vision Workshops (ICCVW)*, 2019, pp. 321–328. [Online]. Available: https://doi.org/10.1109/ICCVW.2019.00041

[5] H. Proença and J. C. Neves, "Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2078–2088, 2019. [Online]. Available: https://doi.org/10.1109/TIFS.2019.2892728

[6] K. R. Radhika, S. Gupta, M. J. Oli, D. Sharma, H. Vinutha, and H. N. Shenoy, "Pancreatic tumor segmentation using conditional diffusion and graph regularization on abdominal computed tomography images," *Journal of Intelligent Decision Making and Information Science*, vol. 3, no. 1s, pp. 1123–1143, 2026. doi: 10.59543/jidmis.v3.489

[7] T.-Y. Lin *et al.*, "Focal Loss for Dense Object Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, 2020.

[8] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020. [Online]. Available: https://arxiv.org/abs/2004. 10934

[9] G. Jocher *et al.*, "YOLOv5," *GitHub Repository and Documentation*, Ultralytics, 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[10] Z. Zhao, H. Zhang, and X. Li, "Real-Time Eye Tracking Using Deep Learning for Human-Computer Interaction," *IEEE Access*, vol. 8, pp. 156789–156799, 2020.

[11] P. P. Basavaraju, S. Gupta, S. Khullar, H. R. Seema, G. T. Varma, H. N. Shenoy, and A. M. Amaresh, "MPC-Net: Detection of chronic kidney disease using missingness-pattern conditioned deep learning," *Natural Resources for Human Health*, vol. 6, no. 6s, pp. 1333–1349, 2026.

[12] P. Rot, Z. Emeršiç, V. Struc, and P. Peer, "Deep Multi-Class Eye Segmentation for Ocular Biometrics," in *IEEE International Workshop on Biometrics and Forensics (IWBF)*, 2020, pp. 1–6.

[13] M. Zhang, Y. Liu, and Z. Sun, "Fast Iris Segmentation Using Convolutional Neural Networks," *Pattern Recognition Letters*, vol. 138, pp. 567–573, 2020.

[14] N. Bouguila, H. Gaceb, and M. Deriche, "Bayesian Learning for Robust Eye Detection in Unconstrained Environments," *Computer Vision and Image Understanding*, vol. 168, pp. 87–101, 2018. [Online]. Available: https://doi.org/10.1016/j.cviu.2017.12.006

[15] F. H. Dahri *et al.*, "Real-Time Detection of UAV Small Target's Infrared Imaging via OSTDYOLOv8," in *2024 IEEE URUCON*, 2024. doi: 10.1109/URUCON63440.2024.10850067.

[16] A. Shazia, F. H. Dahri, A. Ali, M. Adnan, A. A. Laghari, and T. Nawaz, "Automated Early Diabetic Retinopathy Detection Using a Deep Hybrid Model," *IECE Transactions on Emerging Topics in Artificial Intelligence*, vol. 1, no. 1, pp. 71–83, 2024.

[17] K. T. Rajgopal, A. F. J. Alshammari, K. Rashmi, R. Puneeth, S. Gupta, and Y. Ambika Naik, "An automated healthcare framework for single-lead ECG arrhythmia classification using lightweight CNN-BiLSTM and temporal attention," *Natural Resources for Human Health*, vol. 6, no. 1s, pp. 255–266, 2026.

[18] M. Javed, Z. Zhang, F. H. Dahri, and A. A. Laghari, "Real-Time Deepfake Video Detection Using Eye Movement Analysis with a Hybrid Deep Learning Approach," *Electronics*, vol. 13, no. 15, 2024. doi: 10.3390/electronics13152947.

[19] F. H. Dahri, G. E. M. Abro, N. A. Dahri, A. A. Laghari, and Z. A. Ali, "Advancing Robotic Automation with Custom Sequential Deep CNN-Based Indoor Scene Recognition," *IECE Transactions on Intelligent Systems*, vol. 2, no. 1, pp. 14–26, 2024.

[20] M. Javed, Z. Zhang, F. H. Dahri, A. A. Laghari, M. Krajčík, and A. Almadhor, "Audio–Visual Synchronization and Lip Movement Analysis for Real-Time Deepfake Detection," *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, 2025. doi: 10.1007/s44196-025-00911-7.

[21] M. Javed, Z. Zhang, F. H. Dahri, and T. Kumar, "Enhancing Multimodal Deepfake Detection with Local–Global Feature Integration and Diffusion Models," *Signal, Image and Video Processing*, vol. 19, no. 5, pp. 1–9, 2025.

[22] B. Singh and L. Nguyen, "Comparative Analysis of YOLO and Faster R-CNN for Real-Time Eye Detection," *Computer Vision and Image Understanding*, vol. 206, p. 103123, 2021.

[23] P. K. Sekharamantry, P. MeganaSanthoshi, S. Gupta, M. Akshatha, K. L. Divya, and M. Swathi Pai, "CAF-Net: A convolution and attention fusion network with RR interval context for five-class arrhythmia detection from single-lead ECG," *Natural Resources for Human Health*, vol. 6, no. 2s, pp. 452–465, 2026.

[24] L. Zhao and T. Yu, "Framework for Biometric Iris Recognition in Video Using YOLO," *International Journal of Computer Vision*, vol. 129, no. 3, pp. 453–467, 2021. [Online]. Available: https://doi.org/ 10.1007/s11263-020-01389-8

[25] P. Taylor and C. Morris, "Deep Learning for Iris Recognition: A Review," *Journal of Biometrics and Pattern Recognition*, vol. 37, no. 5, pp. 1123–1138, 2021.

[26] M. Chen, S. Gao, and X. Zhang, "Fast and Accurate Pupil Localization in Natural Scenes," *Image and Vision Computing*, vol. 105, p. 104056, 2021. [Online]. Available: https://doi.org/10.1016/j. imavis.2020.104056

[27] S. Patel and J. Brown, "Advancements in Iris Detection Using Convolutional Neural Networks," *Neural Processing Letters*, vol. 53, no. 3, pp. 502–517, 2021. [Online]. Available: https://doi.org/10. 1007/s11063-020-10394-8

[28] J. Daugman, "How Iris Recognition Works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004. [Online]. Available: https://doi.org/10.1109/TCSVT. 2003.818350

[29] S. Gupta, U. N. Ranjitha, J. G. S. Banu, K. L. Divya, T. R. Kulkarni, and S. Mannikeri, "Adaptive multi-band dynamic functional-connectivity graph attention network for Parkinson's disease classification from resting-state EEG," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 18, no. 11s, pp. 840–858, 2026. doi: 10.70917/ijcisim-2026-3819

[30] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image Understanding for Iris Biometrics: A Survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008. [Online]. Available: https://doi.org/10.1016/j.cviu.2007.08.001

[31] Q. Xu, X. Zhang, Y. Wang, X. Chen, and L. Zhou, "Robust Iris-Localization Algorithm in NonCooperative Environments Based on the Improved YOLO v4 Model," *Sensors*, vol. 22, no. 24, pp. 1–18, Dec. 2022.

[32] Y. Zhang, H. Wang, L. Wang, and T. Li, "Deep-Learning-Based Pupil Center Detection and Tracking Technology for Visible-Light Wearable Gaze Tracking Devices," *Applied Sciences*, vol. 11, no. 2, pp. 1–14, Jan. 2021.

[33] M. Frasca and D. La Torre, "An Extended Daugman's Algorithm for Iris with Eye Pathology Recognition," *SSRN Electronic Journal*, Feb. 2024. [Online]. Available: https://ssrn.com/abstract= 4694096. doi: 10.2139/ssrn.4694096

[34] F. J. Paulín-Martínez, A. Lara-Guevara, R. M. Romero-González, and H. Jiménez-Hernández, "Implementation of the Hough Transform for Iris Detection and Segmentation," *Advances in Molecular Imaging*, vol. 9, no. 1, pp. 6–18, 2019. doi: 10.4236/ami.2019.91002

[35] R. Gao and T. Bourlai, "On Designing a SwinIris Transformer Based Iris Recognition System," *IEEE Access*, pp. 1–1, 2024. doi: 10.1109/ACCESS.2024.3369035

[36] V. Tikhonova and E. Pavelyeva, "Hybrid Approach to Iris Segmentation Using Modified Daugman's Method, Dual U-Net CNNs and Principal Curvature Analysis," in *Proc. 30th Int. Conf. Computer Graphics and Machine Vision (GraphiCon 2020)*, pp. 1–10, 2020. doi: 10.51130/graphicon-2020-23-31

[37] L. K. Singh, H. Garg, and M. Khanna, "An artificial intelligence-based smart system for early glaucoma recognition using OCT images," *International Journal of E-Health and Medical Communications (IJEHMC)*, vol. 12, no. 4, pp. 32–59, 2021.