Sensors and Materials, Vol. 37, No. 12 (2025) 5717–5731
MYU Tokyo

S & M 4274

Deep-learning-driven Online Tool for Automated Strabismus Diagnosis

San-Yuan Wang,[1] Geng-Bin Liu,[1] Ming-Chih Chien,[1]
Yu-Hung Lai,[2,3] and Rong-Ching Wu[4*]

[1]Department of Information Engineering, I-Shou University,
No. 1, Sec. 1, Syuecheng Rd., Dashu Dist., Kaohsiung City 84001, Taiwan, R.O.C.
[2]Department of Ophthalmology, Kaohsiung Medical University Hospital, Kaohsiung Medical University,
Kaohsiung 80756, Taiwan, R.O.C.
[3]Department of Ophthalmology, School of Medicine, College of Medicine, Kaohsiung Medical University,
Kaohsiung 80708, Taiwan
[4]Department of Electrical Engineering, I-Shou University,
No. 1, Sec. 1, Syuecheng Rd., Dashu Dist., Kaohsiung City 84001, Taiwan, R.O.C.

(Received November 18, 2025; accepted December 16, 2025)

Keywords: strabismus diagnosis, deep learning, YOLOv8n, online tool

Strabismus, characterized by ocular misalignment and loss of binocular fixation, can result in impaired eye movement, blurred vision, asthenopia, abnormal head posture, reduced stereopsis, and, if untreated in early childhood, amblyopia. Conventional diagnosis depends on in-person ophthalmologic examination and specialized instruments, which limits early detection in resource-constrained settings. In this study, we present an image-based vision-sensing system that transforms a commodity camera into a low-cost ophthalmic screening sensor for automated strabismus detection. Facial images captured by the RGB camera are processed using You Only Look Once v8 Nano (YOLOv8n) deep learning model to localize ocular regions and analyze ocular alignment patterns that reflect the presence of strabismus. The proposed online platform provides a simple web-based user interface and achieves a diagnostic accuracy of 95.48%, demonstrating that consumer-grade imaging sensors, when combined with advanced deep learning algorithms, can function as effective medical screening devices. The results of this work contribute to the advancement of sensor applications in ophthalmology by enabling scalable, noncontact, and remote vision screening, thereby broadening the practical use of imaging sensors and sensor-enabled materials in telemedicine and community-based eye health programs.

1. Introduction

The eyes serve as one of the most vital sensory organs in humans. Normal visual development plays a pivotal role in early childhood learning, exploration, and cognitive growth.(1) If visual abnormalities are not identified and treated during the critical period of visual maturation, they may lead to irreversible visual deficits. Clinical observations have demonstrated that the visual

*Corresponding author: e-mail: rcwu@isu.edu.tw
https://doi.org/10.18494/SAM6055

ISSN 0914-4935
https://myukk.org/

system of young children retains high plasticity; timely intervention during early childhood can yield excellent therapeutic outcomes. However, if the opportunity for intervention is missed before approximately six years of age, visual system maturation may limit the effectiveness of subsequent treatment.(2)

Strabismus, characterized by ocular misalignment, prevents the two eyes from maintaining proper visual axes and disrupts binocular fusion. As shown in Fig. 1, there are four types of strabismus: *esotropic*, *exotropic*, *hypertropic*, and *hypotropic*. To avoid diplopia, the brain suppresses the input from the deviating eye, leading to amblyopia.(3) In addition to reduced visual acuity, affected individuals often lose stereopsis, binocular coordination, and spatial depth perception. When strabismus develops during infancy and remains untreated, it can cause long-term degeneration or even irreversible vision loss in the amblyopic eye.(4) Older children may experience visual fatigue, headaches, or compensatory head-tilt postures to minimize double vision.(5) These conditions substantially affect both visual performance and quality of life.

1.1 Clinical diagnosis and current limitations

Conventional diagnostic procedures for strabismus include photo-screening devices for infants and toddlers, which provide rapid and noninvasive identification of ocular deviations and refractive errors. Pediatric ophthalmologists also employ prism cover tests, ophthalmoscopy, retinoscopy, and handheld autorefractors to screen for anisometropia, strabismus, and other visual anomalies.(6) Diagnosis is typically made by examining whether the iris and corneal reflex are symmetrically aligned along the visual axis.

Early detection and treatment are essential for preventing permanent amblyopia and ensuring normal binocular vision development. Failing to diagnose and correct strabismus or amblyopia within the sensitive developmental window can lead to long-term, irreversible impairment that extends beyond visual acuity to affect cognitive and motor coordination. As such, comprehensive

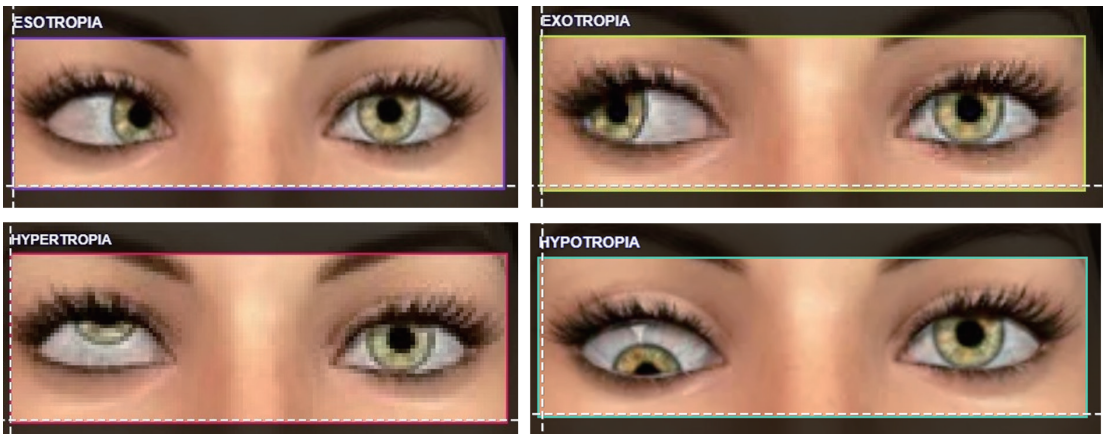


Fig. 1. (Color online) Four types of strabismus.

early screening and timely intervention are critical for preserving visual health and supporting learning development in preschool-aged children.

Ophthalmologic guidelines recommend that children undergo eye examinations between 0 and 2 years of age, particularly when warning signs such as the following are present:[7] (1) obvious ocular deviation or misalignment, (2) habitual head tilt, (3) unilateral squinting under bright light, (4) frequent falls or poor spatial orientation, and (5) other abnormal visual behaviors. The symptoms mentioned are too obscure for parents to distinguish and to correlate to strabismus. Despite its clinical importance, strabismus often goes unnoticed by parents, especially in initial stages or in communities with limited access to ophthalmic specialists. Consequently, the optimal treatment period is missed for many children, particularly in under-resourced or rural regions. If potential cases can be detected earlier, both diagnostic accuracy and surgical success rates can be improved.

1.2 Motivation for automated and accessible screening

At present, most strabismus diagnoses depend on in-person evaluation by ophthalmologists. There is a lack of accessible, automated, and self-administered screening tools that can facilitate early detection outside clinical settings. However, early recognition and intervention are critical for preventing amblyopia and ensuring successful treatment outcomes.

The increasing ubiquity of smartphones and internet connectivity offers a unique opportunity to bridge this diagnostic gap. By combining advances in computer vision and deep learning, image-based analysis of uploaded facial photographs can enable scalable, low-cost screening that is both rapid and objective. Such systems can benefit pediatric populations, individuals with limited mobility, and children living in remote or medically underserved areas.

In this study, we aim to develop an image-based automated strabismus detection system that leverages deep learning and large-scale data analysis. The proposed framework seeks to identify ocular misalignment from user-submitted images, providing immediate feedback through an intuitive web interface. By integrating AI with accessible web technologies, our objective is that the system will democratize visual health assessment and assist clinicians by prescreening potential cases of strabismus.

1.3 Relative works

For computer-based detection, a prevalent strategy employs MediaPipe to localize periocular landmarks,[8] combined with a deep-learning model to delineate the iris contour, identify the specular highlight (glint), and estimate the pupil/iris center to assess ocular deviation. A recent system proposed by Şükrü *et al.*[8] operationalizes the Hirschberg method by quantifying the horizontal and vertical distances between the corneal reflex and the pupil center to estimate the strabismic angle; a centrally located reflex indicates orthotropia, whereas offsets correspond to horizontal and/or vertical deviations. MediaPipe provides 478 facial landmarks, enabling the precise localization of the eyes and iris; once localized, the iris region is further refined. The iris boundary is approximated via a minimum-enclosing-circle procedure to mark the iris center,

and a sequence of masking and binarization steps is applied to detect the glint position. The pupil boundary is refined in the hue, saturation, value (HSV) color space to enhance separation between the specular reflection and the pupil contour.

An image-processing-based automated strabismus screening method has been proposed by Huang *et al.*[9] The facial regions are first detected using Dlib's convolutional neural network (CNN)-based face detector, and then ocular regions are localized via the 68-point facial landmark model. Otsu's automatic thresholding[9] is applied to image preprocessing to determine an optimal binarization threshold, suppressing skin and scleral (ocular white) interference while preserving the iris. In parallel, the HSV color space is used to separate the iris from background artifacts (e.g., eyelash shadows); the HSV-based segmentation is then fused with the Otsu result to generate a refined processed image. To estimate the iris/pupil center, boundary points are sampled along the limbus, and the center coordinates are computed by the least squares method. When the ratio of distances from the pupil center to the inner and outer canthi approaches one, the eyes are considered symmetric; conversely, ratios exceeding a predefined tolerance suggest potential strabismus.

Early geometric and Hough-transform approaches[10,11] established proof-of-concept systems capable of detecting horizontal misalignment in controlled settings. Machine learning classifiers[12,13] improved detection accuracy but remained constrained by feature engineering. The advent of deep learning[14–18] enabled automated, data-driven feature extraction, achieving performance levels approaching clinical standards.

Recent works emphasize real-world deployment by integrating smartphone imaging,[11,16] VR calibration,[19,20] and eye tracking[21] to enhance measurement accuracy and accessibility. Moreover, emerging frameworks[22,23] highlight telemedicine integration, reflecting the field's transition from algorithmic experimentation to holistic healthcare applications.

Despite these advances, several critical issues persist. First, dataset limitations remain a primary bottleneck: most models are trained on small, geographically localized datasets with limited demographic diversity, reducing generalizability. Second, illumination and pose variation continue to challenge image-based systems. Robust alignment and normalization strategies are required to ensure consistent gaze feature extraction. Third, model interpretability—a crucial factor for clinical acceptance—has been insufficiently addressed; few studies provide visual explanations or saliency analyses linking image features to diagnostic outcomes. Finally, cross-platform validation (e.g., testing on data from different camera types or acquisition protocols) remains rare, limiting real-world applicability. In addition, current methods generally output binary classification (strabismus vs normal) without quantifying deviation angle or severity level, which restricts their clinical interpretability. A summary of the related works is provided in Table 1.

Given these limitations, the aim of the present research is to develop a robust image-based strabismus detection framework capable of operating under realistic conditions, ensuring generalization across demographic and device variations. The system emphasizes model explainability, lightweight architecture, and potential scalability for mobile or telemedicine applications.

Table 1
Summary of the representative studies discussed in this review, categorized by their methodological approach, data modality, and main contribution.

Year	Study	Approach	Modality	Key contribution
2018	Gupta[11]	Classical vision + ML	Smartphone images	Mobile heterotropia diagnosis
2019	Simões *et al.*[15]	U-Net + ResNet	Ocular photographs	Automatic alignment evaluation
2020	Zolkifli & Nazari[10]	Hough transform	Image geometry	Iris tracing and angle estimation
2020	Mehringer *et al.*[19]	Image-based VR	Virtual reality	3D strabismus measurement
2021	Kim *et al.*[14]	CNN classification	Facial photographs	End-to-end detection
2021	Mehringer *et al.*[20]	VR + Eye tracking	Dynamic gaze	Hess screen modernization
2021	Huang *et al.*[9]	CNN classification	Facial photographs	End-to-end detection
2023	Daher *et al.*[22]	AI diagnostic platform	Multimodal	Integrated detection and rehab
2023	Ebrahimzadeh *et al.*[21]	Ensemble learning	Eye tracking	Severity quantification
2023	Rohismadi *et al.*[12]	ML classifiers	Binocular vision	Automated classification
2023	Şükrü *et al.*[8]	Deep learning model	Hirschberg method	Computer-aided classification
2024	Chang *et al.*[16]	Deep CNN	Smartphone photos	Pediatric large-scale screening

1.4 Research contributions

We propose a web-based online tool for strabismus detection. There are three key contributions.

(1) Diagnostic accuracy: a robust deep learning model is adopted for visual recognition that accurately distinguishes between normal and strabismic eye alignment, thereby enhancing early detection reliability.
(2) User-friendly interface: an intuitive web-based platform suitable for users of all ages is implemented, enabling individuals, parents, or school personnel to perform preliminary screening conveniently at any time.
(3) Real-time feedback: near-instantaneous diagnostic output that provides preliminary results for subsequent professional evaluation is realized, thereby supporting early medical consultation and intervention.

The integration of the YOLOv8n deep learning technique with web-based image acquisition provides an innovative solution to the long-standing problem of accessibility in pediatric vision screening. Unlike traditional manual examinations, automated image-based systems can perform consistent, quantitative analysis without relying on specialist availability. In public-health contexts, such systems can increase screening coverage, reduce diagnostic delays, and facilitate remote triage for ophthalmologists. Moreover, by enabling parental participation through online tools, awareness and early referral rates for childhood strabismus are expected to improve substantially. From a technological perspective, this research also contributes to the growing field of AI-assisted ophthalmology, as we demonstrate how modern computer-vision pipelines—such as CNNs and facial landmark analysis—can be adapted for clinical decision support. The broader impact lies in advancing accessible, data-driven eye-health monitoring aligned with global digital-health initiatives.

1.5 Organization of the paper

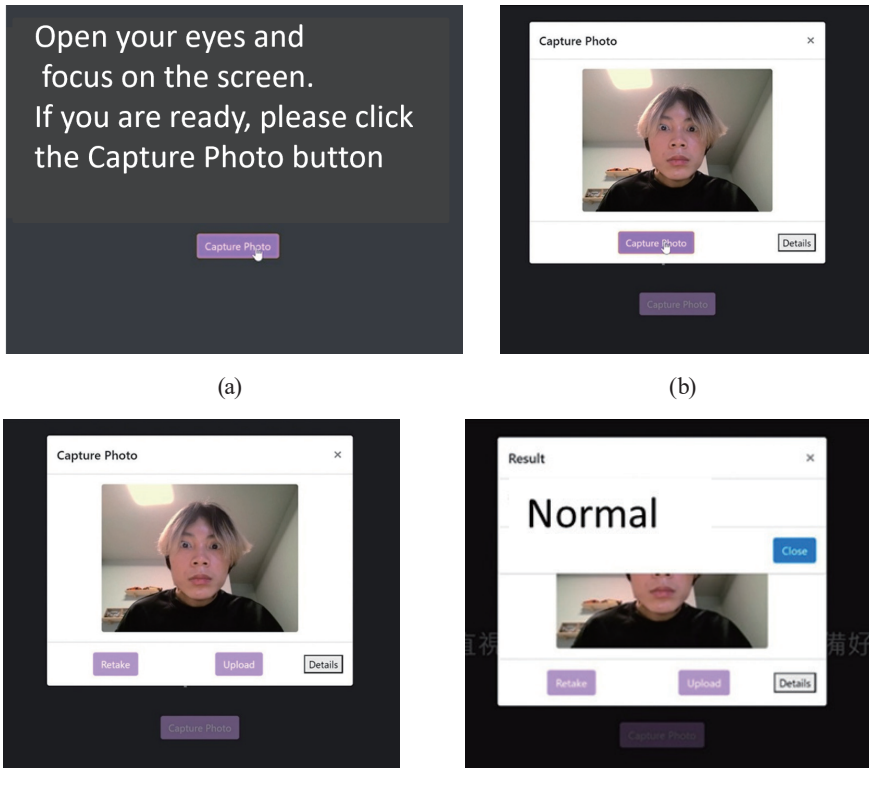
The remainder of this paper is organized as follows. In Sect. 2, we describe the proposed web-based strabismus detection system. The strabismus detection model—the core of our web-based strabismus detection system—is presented in Sect. 3. Finally, in Sect. 4, we conclude the paper and directions for future work are outlined.

2. Web-based Strabismus Detection System

2.1 User interface design

To enable users to conveniently perform self-screening for potential strabismus, a web-based interface is developed emphasizing accessibility and ease of use. The interface is implemented in HyperText Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript, with a focus on intuitive operation and clear workflow.

The HTML structure provides a minimal yet functional layout. The initial screen is shown in Fig. 2(a). There are prompt messages and a Capture Photo button. There is a preview screen [see Fig. 2(b)] for users to check the photo status. When a user clicks the Capture Photo button, the



(a) (b)

(c) (d)

Fig. 2. (Color online) Client site user interface. (a) Initial screen of user interface. (b) Preview screen. (c) Screen after capturing. (d) Feedback result.

user's face is shown on the screen. Retake and Upload buttons then appear, as can be seen in Fig. 2(c). The user uploads the photo and after a wait of a few seconds, informative text at the top of the page summarizes the types of strabismus to assist user interpretation. Figure 2(d) illustrates the Normal state on the feedback result page.

The CSS layer enhances visual clarity and organization, ensuring a clean and professional appearance suitable for both desktop and mobile browsers. The JavaScript back-end assigns functionality to all buttons and coordinates communication with the server. Using Node.js, the client issues asynchronous HyperText Transfer Protocol (HTTP) requests that trigger camera operations via commands such as "command_to_start-camera_and_flash." When a button is clicked, a corresponding function is executed.

Captured images are automatically stored and transmitted to the server for processing. Once received, the server loads the trained strabismus detection model and performs inference on the uploaded image. The prediction result is then returned to the web application, where conditional logic determines the specific type of strabismus detected. The diagnostic outcome—both as an annotated image and textual description—is displayed in the designated result area. After completion, the interface can be reset by clicking the Initialize button to allow subsequent screening sessions.

2.2 System architecture

On the client side, the user captures a frontal facial photograph using their device's built-in camera. Prior to image acquisition, the system verifies whether the user's face is within an acceptable capture distance, as improper distance may affect detection accuracy. When optimal alignment and distance are achieved, the Capture Photo button becomes active, allowing the user to take and upload the facial image. When the client receives a "did not recognize the face" warning signal from the server side, the client system verifies whether the user's face lies within an acceptable capture distance. The left-and right-hand sides of Fig. 3 illustrate the client-side and server-side workflows, respectively.

On the server side, two input modes are supported.

(1) For full-face images, the system employs MediaPipe to automatically detect the eye region and crop it for analysis.
(2) For precropped eye images, the model goes directly to preprocessing without additional localization.

The preprocessing pipeline begins with automatic face and eye detection using MediaPipe. If the face and eye are not detected, a warning signal is sent back to the client side and the server side goes back to the demon process and waits for uploaded images Once the ocular regions are identified, each image is resized to a uniform resolution of 720 pixels in height, maintaining the original aspect ratio to prevent geometric distortion. The pixel intensity is normalized to the [0, 1] range, and color channels are standardized according to the ImageNet mean and variance parameters. This ensures compatibility with pretrained deep learning backbones used in the feature extraction process.

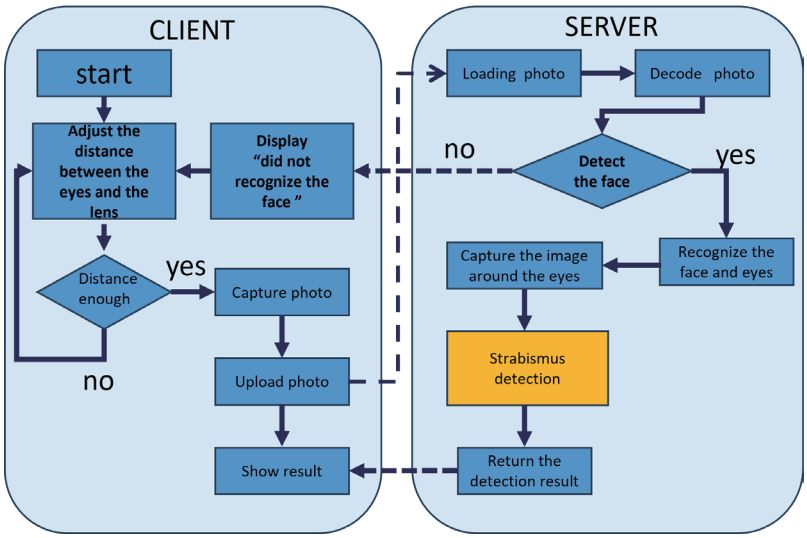


Fig. 3. (Color online) Overall workflow architecture of the system.

The back-end model of the proposed web-based strabismus detection system was implemented using a YOLOv8n framework improved for lightweight inference on server environments. The model architecture integrates multiple convolutional and pooling layers designed to capture both global facial symmetry and localized ocular deviations, enabling precise classification between normal alignment and various strabismus types.

All input images are standardized to a length of 720 pixels, with the width scaled proportionally to preserve the aspect ratio, ensuring consistency during model inference. For the classification task, the final dense layer outputs probability scores corresponding to each diagnostic category: *normal*, *esotropia*, *exotropia*, *hypertropia,* and *hypotropia*. A softmax activation function converts these scores into interpretable likelihoods. The category with the highest probability is selected as the predicted class. Additionally, the model provides a confidence score that quantifies prediction certainty, which is later displayed alongside the diagnostic text on the web interface.

The YOLOv8n-based detection model (details in Sect. 3) then analyzes the processed image to judge the presence and category of strabismus. Finally, the diagnostic result is transmitted back to the client-side web page for immediate visualization and feedback to the user.

2.3 Server system implementation and deployment

The trained model is deployed on a Node.js-integrated Python back-end server, where inference is executed through a lightweight RESTful application programming interface (API). When an image is uploaded from the client-side web interface, the API processes the request as follows.

(1) Input acquisition: The uploaded image is received as a base64-encoded object via the Fetch or XMLHttpRequest protocol.

(2) Preprocessing: The server decodes the image, performs cropping and normalization, and prepares it for inference.
(3) Strabismus Detection Model inference: The image tensor is passed to the YOLOv8n model for forward propagation, yielding the classification and confidence scores.
(4) Postprocessing: The prediction result is converted into a human-readable message indicating the type of strabismus (if detected).
(5) Response transmission: The annotated image and textual result are returned to the client-side interface for visualization and record display.

To ensure real-time feedback, the server utilizes asynchronous request handling and GPU-accelerated computation when available. The entire inference process—from image upload to result visualization—is typically completed within a few seconds on a standard server configuration (Intel i7 CPU, 16 GB of RAM, NVIDIA RTX 3060 GPU). This enables efficient screening even under concurrent user sessions. The system's modular design also supports continuous model updating. New labeled data can be incorporated into the training set to fine-tune model parameters, thereby improving accuracy over time.

3. Strabismus Detection Model

In this section, we describe the development of the proposed YOLOv8n-based strabismus detection model, as illustrated in Fig. 4. The MIT Strabismus Dataset[24] was adopted as the primary data source. Each image was manually annotated with its corresponding strabismus

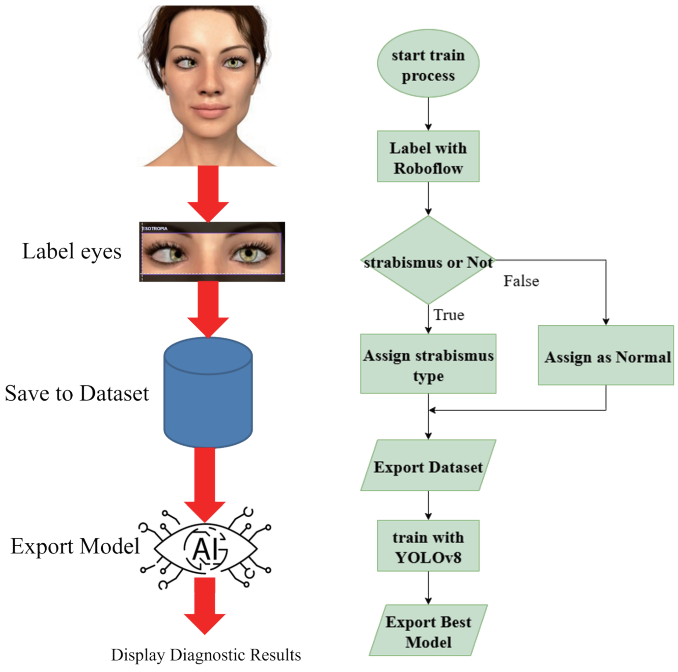


Fig. 4. (Color online) Strabismus detection model procedure.

category using the Roboflow platform.[25] The labeled dataset was then partitioned into training, validation, and test sets for model optimization.

3.1 Data preprocessing

In this study, the MIT Strabismus Dataset[24] was used, comprising 509 images across five ocular categories: normal, esotropia, exotropia, hypertropia, and hypotropia, with 110, 100, 104, 102, and 93 images, respectively. To enhance model performance, all images were manually labeled on the Roboflow platform[25] in an object-detection format. Each image contained one or more bounding boxes highlighting the abnormal eye region, each assigned to the corresponding class label.

After annotation, image quality was quantitatively assessed by the Laplacian variance method to compute a sharpness score, where lower values indicate greater blurriness. For each category, the 10 images with the lowest sharpness scores (50 blurred samples in total) were identified and moved to a separate directory. These severely blurred images were excluded from the main dataset, resulting in 100 normal, 90 esotropic, 94 exotropic, 92 hypertropic, and 83 hypotropic images used as the primary training set.

The annotated images were then automatically converted by Roboflow into a YOLOv8-compatible format, including the corresponding "data.yaml" configuration file. The dataset was split into training, validation, and testing subsets at 80, 10, and 10%, respectively, while maintaining balanced representation across all classes.

3.2 Model training, validation, and testing

The YOLOv8 nano (YOLOv8n) model provided by Ultralytics was selected as the core detector for five-class training because of its lightweight architecture and strong real-time performance. The input image resolution was fixed at 240 × 720 pixels, and the network was fine-tuned for 50 epochs starting from the pretrained yolov8n.pt weights.

The training process was conducted in 11 incremental stages. In the initial stage, only clear images from the main dataset were used. Starting from the second stage, in each iteration, one clear image per class was randomly replaced with a blurred image, thereby introducing five blurred samples per stage. This procedure was repeated until all 50 blurred images had been incorporated. The progressive introduction of degraded samples enabled the systematic evaluation of how image clarity affects model performance and robustness.

After training, the optimized model weights were saved in the file best.pt and used for inference and testing. During inference, the system automatically performed detection on all images in the designated test directory and generated both annotated result images and prediction metadata (class label and confidence score). The inference workflow was as follows.

(1) Load the trained YOLOv8n model weights from best.pt.
(2) Perform inference on each test image to obtain all detected bounding boxes and corresponding class probabilities.
(3) If multiple detections occurred, retain the bounding box with the highest confidence score.

(4) Save the annotated output image to the results directory.
(5) Record the image filename, predicted class ID, and confidence value in a spreadsheet for subsequent quantitative analysis.

3.3 Performance evaluation

The confusion matrix (Table 2) summarizes the classification outcomes for all 509 images. Diagonal cells represent correctly classified samples, where larger values indicate higher accuracy, while off-diagonal elements correspond to misclassifications between neighboring categories or the background class, reflecting ambiguity in subtle boundary cases. Despite the presence of such challenging instances, the trained YOLOv8n model achieved an overall detection accuracy of 95.48%, demonstrating excellent capability in identifying and discriminating among the five types of strabismus from ocular images.

3.4 Model comparison and discussion

3.4.1 Positioning among image-based methods

Classical CNN classifiers for strabismus typically operate on cropped eye or face images and output a single label without spatial localization.[14] Segmentation-oriented pipelines (e.g., U-Net/ResNet) localize fine ocular structures but often require pixel-level masks and heavier annotation cost.[15] In contrast, our approach formulates strabismus analysis as object detection with five classes (normal, esotropia, exotropia, hypertropia, hypotropia), producing bounding boxes around abnormal ocular regions and a class label in one pass. This affords (i) spatial interpretability, (ii) compatibility with mixed full-face/eye-crop inputs, and (iii) real-time feasibility on modest hardware, making it well suited to web deployment.

Table 2
(Color online) Confusion matrix of strabismus detection model.

predicted/true	ESOTROPIA	EXOTROPIA	HYPERTROPIA	HYPOTROPIA	NORMAL	Background
ESOTROPIA	93	0	1	1	0	10
EXOTROPIA	0	97	2	0	1	7
HYPERTROPIA	2	4	97	1	0	0
HYPOTROPIA	3	2	0	91	1	3
NORMAL	2	1	2	0	108	6
Background	0	0	0	0	0	74

3.4.2 Data preparation and robustness

A recurring limitation reported in the literature is sensitivity to image blur, illumination, and pose particularly for smartphone photographs used in school/tele-screening scenarios.[16–18] Our pipeline explicitly quantifies blur via Laplacian variance, sets aside the worst 10 images per class (50 total), and then progressively reintroduces blurred samples during training (11 stages), which improves tolerance to degraded inputs. This staged curriculum differs from standard augmentation and helps the detector internalize blur-robust cues without collapsing localization accuracy.

3.4.3 End-to-end detection quality

Compared with pure classifiers, the detection quality of our system must satisfy both localization and classification constraints. Our raining/validation traces show stable convergence of box regression and distribution focal loss, indicating refined boundary placement, while classification loss drops rapidly without overfitting. Our model attains mAP@50 ≈ 0.70 and mAP@50–95 ≈ 0.60, reflecting balanced box-and-label performance; the confusion matrix over all 509 images yields an overall accuracy of 95.48%. These figures are consistent with an application-grade detector for multitype ocular misalignment in mixed-quality images.

3.4.4 Comparison with VR/eye-tracking systems

VR eye-tracking and Hess-screen modernizations can produce precise angle measurements but need specialized hardware (HMD + trackers) and clinic-like calibration.[19–21] Our web-first detector trades absolute angle estimation for ubiquity and low friction, enabling rapid prescreening at scale, which aligns with pediatric school programs and tele-ophthalmology workflows.[16,18] In practice, these modalities can be complementary: web detection for triage and VR/eye-tracking for quantitative follow-up.

3.4.5 Summary

Overall, the proposed YOLOv8n detector achieves application-grade multiclass detection with real-time web deployment, directly addressing accessibility and speed gaps identified in prior work, while leaving clear paths to clinical enrichment (angle estimation, external validation, explainability). All dataset composition, preprocessing, staged training with blurred samples, and metric values summarized here derive from our implementation and results. Through robust data preprocessing, careful augmentation, and staged training incorporating both clear and blurred samples, the model achieved high precision, recall, and overall accuracy. The architecture's lightweight design also enables real-time deployment within the web-based diagnostic platform, providing a practical and scalable solution for automated strabismus screening.

4. Conclusions

Strabismus can lead to disordered ocular motility, blurred vision, asthenopia, compensatory head tilt, reduced stereopsis, and psychosocial burden due to cosmetic concerns. Although timely identification in regions with advanced medical care, such as Taiwan, usually enables effective treatment, access to ophthalmologic examination and specialized diagnostic instruments remains limited in some communities. In this work, we developed a strabismus diagnostic system that uses a conventional RGB camera as an imaging sensor and a YOLOv8n-based deep learning model as the processing core to detect ocular misalignment from facial images. The system achieved an overall accuracy of 95.48%, showing that a low-cost camera can be repurposed as a practical ophthalmic screening sensor when coupled with appropriate sensing algorithms.

By deploying this sensing system on an internet-based platform, we enable online strabismus screening that reduces the time interval between disease onset and clinical intervention, supports remote prescreening prior to hospital visits, and enhances access for residents in medically underserved areas and individuals with limited mobility. From the perspective of *Sensors and Materials*, the results of this study demonstrate a concrete application in which widely available imaging sensors and digital materials (web and mobile interfaces) are integrated into a vision-sensing framework for clinical decision support.

Future work will be focused on integrating the proposed sensing platform with electronic health record systems or tele-ophthalmology platforms, allowing clinicians to remotely review and verify automated screening results. In addition, the entire strabismus detection pipeline can be deployed on smartphones or other mobile devices as an edge-AI implementation, further extending the reach of camera-based medical sensing in everyday environments.

Acknowledgments

We are immensely grateful for the support from the Ministry of Science and Technology through project number NSTC 112-2221-E-214 -002 -MY2.

References

1 E. E. Birch and J. M. Holmes: J. AAPOS **14** (2010) 494. https://doi.org/10.1016/j.jaapos.2010.10.004
2 J. M. Holmes and M. P. Clarke: Lancet **367** (2006) 1343. https://doi.org/10.1016/S0140-6736(06)68581-4
3 G. K. von Noorden and E. C. Campos: Binocular Vision and Ocular Motility: Theory and Management of Strabismus (Mosby, 2002) 6th ed.
4 L. B. Nelson: J. Ophthalmic Nurs. Technol. **2** (1983) 157.
5 C. J. Lyons and S. R. Lamnert: Pediatric Ophthalmology and Strabismus (AAOA, San Francisco, CA, 2022) 6th ed.
6 EyeWiki, Pediatric vision screening: https://eyewiki.org/Pediatric_Vision_Screening (accessed November 2025).
7 S. A. Cotter, L. A. Cyert, J. M. Miller, and G. E. Quinn: Optom. Vision Sci. **92** (2015) 6. https://doi.org/10.1097/OPX.0000000000000429
8 K. Şükrü, S. G. Kobat, and M. Gedikpınar: Photodiagn. Photodyn. Ther. **44** (2023) 103805.
9 X. Huang, S. J. Lee, C. Z. Kim, and S. H. Choi: PLOS ONE **16** (2021) e0255643.
10 N. S. Zolkifli and A. Nazari: IEEE Student Conf. Res. Dev. (SCOReD, 2020) 313–318.

11 H. Gupta: 3rd Int. Conf. Contemporary Computing and Informatics (IC3I, 2018) 160–165.
12 M. A. I. Rohismadi, M. H. Ithnin, A. F. M. Raffei, S. F. Othman, and N. S. A. Zulkifli: IEEE 8th Int. Conf. Software Engineering and Computer Systems (ICSECS, 2023) 487–492.
13 A. Daher and Z. Rammal: 6th Int. Conf. Advances in Biomedical Engineering (ICABME, 2021) 125–128.
14 D. Kim, J. Seo, J. Joo, J. Ha, G. Zhu, and S. C. Kim: Int. Conf. Artificial Intelligence in Information and Communication (ICAIIC, 2021) 588–591.
15 T. O. Simões, J. C. Souza, J. D. S. Almeida, A. C. Silva, and A. C. Paiva: 8th Brazilian Conf. Intelligent Systems (BRACIS, 2019) 239–244.
16 P. Chang, Y. Sun, K. Chuang, L. Yeh, and K. Hsiao: Sci. Rep.**14** (2024) 18277.
17 C. Zheng, Q. Yao, J. Lu, X. Xie, S. Lin, Z. Wang, S. Wang, Z. Fan, and T. Qiao: Transl. Vision Sci. Technol. **10** (2021) 33.
18 J. Lu, Z. Fan, C. Zheng, J. Feng, L. Huang, W. Li, and E. D. Goodman: arXiv preprint arXiv:1809.02940, 2018. Dec.
19 W. A. Mehringer, M. G. Wirth, S. Gradl, L. S. Durner, M. Ring, A. F. Laudanski, B. Eskofier, and G. Michelson: IEEE Int. Symp. Mixed and Augmented Reality Adjunct (ISMAR-Adjunct, 2020) 5–10.
20 W. Mehringer, M. Wirth, F. Risch, D. Roth, G. Michelson, and B. Eskofier: 43rd Annu. Int. Conf. IEEE Engineering in Medicine & Biology Society (EMBC, 2021) 2058–2062.
21 M. Ebrahimzadeh, A. Mohammadpour, and M. Ebrahimzadeh: Biomed. Signal Process. Control. **86** (2023) 104310.
22 A. Daher, M. Ayache, A. J. Zaylaa, M. Z. A. Hadba, Y. Aycha, and H. Hmadeh: 2023 7th Int. Conf. Advances in Biomedical Engineering (ICABME, 2023) 220–224.
23 D. Wu, X. Huang, L. Chen, P. Hou, L. Liu, and G. Yang: Exp. Biol. Med. **249** (2024) 10320. https://doi.org/10.3389/ebm.2024.10320
24 MIT STRABISMUS Dataset: https://www.kaggle.com/datasets/ananthamoorthya/strabismus (accessed on Aug. 28 2025).
25 Roboflow Platform: https://roboflow.com/ (accessed November 2025).

About the Authors

San-Yuan Wang received his B.S. and Ph.D. degrees from the Department of Computer Science and Information Engineering, Tamkang University, Taiwan in 1991, and the National Central University, Taiwan in 1998, respectively. Since 2000 and 2004, he has been an assistant professor and associate professor, respectively, at the Department of Information Engineering, I-Shou University, Taiwan. His research interests include wireless networks, mobile computing, and artificial intelligence. (sywang@isu.edu.tw)

Geng-Bin Liu is currently pursuing his bachelor's degree in the Department of Computer Science and Information Engineering of I-Shou University, Kaohsiung, Taiwan, R.O.C. His research interests include AI and image processing. He is currently engaged in research on a deep-learning-driven online tool for automated strabismus diagnosis and a Ganoderma (Lingzhi) quality inspection system, aiming to apply AI-based techniques to medical and agricultural image analysis. (bsteven.7713553@gmail.com)

Ming-Chih Chien is currently a junior student in the Department of Computer Science and Information Engineering at I-Shou University, Kaohsiung, Taiwan, R.O.C. He has been serving as a research student at I-Shou University, where he has been involved, since 2024, in the design and implementation of visual recognition systems and web server integration. (cleochien0328@gmail.com)

Yu-Hung Lai, M.D., Ph.D., is an attending ophthalmologist at Kaohsiung Medical University Hospital and an adjunct assistant professor at Kaohsiung Medical University. He specializes in pediatric ophthalmology, including preschool vision screening, amblyopia management, and strabismus surgery, with a growing focus on ocular genetics and retinoblastoma. Dr. Lai completed an observational fellowship in pediatric ophthalmology at the University of California, San Francisco (2003–2004) and a fellowship in ocular genetics at Wills Eye Hospital, Philadelphia (2014–2015). Upon returning to Taiwan, he established the ocular genetics clinic at KMU and leads clinical and translational research on pediatric vision and hereditary eye disease. He received his Ph.D. degree from the Graduate Institute of Medicine, Kaohsiung Medical University, in 2018. His current work is integrating population-based screening with genetic evaluation to improve early detection and outcomes. (dyuhung.lai@gmail.com)

Rong-Ching Wu graduated from National Taiwan Institute of Technology, Taipei, Taiwan, in 1990. He received his M.Sc. degree in 1994 and Ph.D. degree in 2001 from National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C. From 1991 to 2001, he was an electrical engineer in Taiwan Power Company. His research interests are in embedded systems, bioengineering, and sensors. Currently, he is a professor at I-Shou University, Kaohsiung, Taiwan, R.O.C. (rcwu@isu.edu.tw)