Grenze International Journal of Engineering and Technology, June Issue

Squint Eye Detection using Machine Learning and Computer Vision

Neelam Chandolikar[1], Hardik Chaudhari[2], Aditya Chavan[3], Pushkar Chandratre[4] and Sonit Chaudhari[5]
[1-5]Vishwakarma Institute of Technology/ Department of Engineering Science and Humanities (DESH), Pune, Maharashtra India
Email: neelam.chandolikar@vit.edu, hardik.1251010715@vit.edu, aditya.1251010398@vit.edu, pushkar.1251010364@vit.edu, sonit.1251010446@vit.edu

***Abstract*— Strabismus, commonly referred to as squint eye, is an ocular disease that entails both eyes not being able to look at the same point at the same time; one of the eyes moves either inwards, outwards, upwards or downwards. If strabismus is not detected early enough, it results in amblyopia, which causes permanent vision loss and cannot be restored after the maturity of the vision system. Existing methods to diagnose the disease involve ophthalmologists, thus requiring expensive tools, making it unaffordable in rural India. This research article presents an automated near real-time squint eye detection system based on MediaPipe Face Mesh, OpenCV, Python programming language, and Random Forest algorithm, trained with 16,654 eye images covering five categories – normal, esotropia, exotropia, hypertropia, and hypotropia. The method uses iris landmarks to calculate deviations of angle ratios between eye corners and irises, and makes use of hybrid decision-making with geometry and machine learning. The model attained accuracy of 91.4%, weighted F1-score of 91% and macro average F1-score of 83% on 3,331 images of test dataset. It works on regular laptops without GPU, produces PDF reports and is integrated into a web application.**

***Index Terms*— Computer Vision, Image Processing, Machine Learning, MediaPipe, OpenCV, Python, Random Forest, Squint Eye Detection, Strabismus.**

I. INTRODUCTION

Strabismus is a disorder where both eyes do not point towards the same direction simultaneously. During an instance when one eye is looking at a particular object, the other eye may look inwardly (esotropia), outwardly (exotropia), upwards (hypertropia), or downwards (hypotropia). As a result of this abnormality, there are two sets of images sent to the brain at the same time, thus making it impossible for the brain to create a three-dimensional picture. Over time, the brain will adapt by suppressing the input from the misaligned eye, causing amblyopia, a condition that weakens vision in that eye. It is important to detect strabismus early because treating it before the child turns eight years old is more effective and less likely to lead to permanent blindness. While being such an important procedure, access to screening for strabismus is quite unequal across the territory of India. While urban settings have opportunities for clinical diagnosis via various tests, such as the cover-uncover test, as well as the Hirschberg corneal light reflex test, there is still a lack of opportunities in rural and underserved areas. Hence, detection takes place when symptoms are visible, causing irreversible consequences for vision.

Grenze ID: 01.GIJET.12.2.663_26

In the current study, an efficient approach in which the process of squint detection is achieved in two steps has been suggested by utilizing geometric calculations and machine learning algorithms without affecting the efficiency of the entire procedure in any manner. Under the suggested method, a live stream of video from the webcam or an image is taken into account as the input data, and the MediaPipe Face Mesh algorithm is employed to calculate the location of the iris landmarks and eye deviation angle. Finally, the output is produced through the combination of results obtained from both geometric calculations and machine learning models.

II. Literature Review

Kathpal et al. [1] developed iChat an eye-based chatting application for differently abled users using OpenCV and facial landmarks numbering 68, that identifies eye blinking and eye gaze with an average accuracy rate of 90%, indicating the applicability of facial landmarks in real-time eye tracking.

Kaur and Ponnala [2] proposed a real-time approach to driver's drowsiness detection using SVM with HOG descriptor for face detection and eye aspect ratio calculation with 92.85% accuracy, establishing the effectiveness of SVM-based classification of eye features for recognizing eye disorders.

Mady et al. [3] proposed a system for face recognition based on Random Forest classifier with LBP and HOG features resulting in 97.6% accuracy.

Rohismadi et al. [4] used a computer algorithm with case-based reasoning to automatically classify the strabismus through Binocular Vision Management System that had 91.8% accuracy, 89.29% precision, 92.59% recall, and 90.91% F1 score, which shows that the machine learning technology can automatically classify strabismus very well.

Gupta et al. [5], four types of deep learning models were implemented in the experiment including EfficientNet-B7, ResNet-50, VGGNet-19, and AlexNet to classify five types of strabismus, among which EfficientNet-B7 had the highest accuracy of 84%. Deep learning is feasible yet computationally intensive.

Wang & Jin [6] developed a smartphone-based system for the digital reproduction of an alternate cover test using the technology of deep learning to monitor eye movements of individuals while covering one eye at a time; however, the procedure demands the cooperation of subjects.

Zhao et al. [7] developed eye tracking device for the detection of intermittent strabismus using Random Forest technique was successful in providing accurate results but needed costly custom infrared equipment.

Ancona et al. [8] created an automatic algorithm for Hirschberg test using the U-Net neural network method provided the measurement of iris displacement in the process of strabismus diagnosis but suffered from erroneous measurements due to corneal reflection sensitivity.

Lou et al. [9] used the two types of deep learning architecture for strabismus diagnosis, and the main reason for which the adoption of deep learning techniques into clinical practice is hindered is that deep learning is difficult to interpret.

Lu and Fan [10] used Deep learning algorithms, which were trained using more than 5,000 images from smartphones for telemedicine-based detection of strabismus, were hindered by varying image resolutions from different cameras, making them inaccurate.

Deep learning techniques provide very high precision results; however, they suffer from the necessity of having huge data sets, computationally intensive operations, and lack of interpretability, whereas geometric techniques have the advantage of being light and easily interpretable. This research work seeks to combine the benefits of geometric deviation analysis and the Random Forest classification algorithm.

Though several deep learning algorithms have shown good results in detecting strabismus, most of them use heavy computation and require a lot of data and specific hardware. The existing geometrical approach provides a good level of interpretability, but it might not perform well in borderline situations. There is still a necessity to develop an efficient, yet interpretable model that will utilize geometrical alignment and classification using machine learning techniques.

III. Objective

The specific objectives of the research included the following. To begin with, the objective was to create squint eye detection technology, which is based on computer vision principles and does not rely on deep learning architectures or dedicated components to be feasible within the context of educational and community-based testing, which can require considerable expenditures to implement. The next objective implied the creation of the software able to provide real-time eye alignment evaluation using OpenCV and MediaPipe Face Mesh algorithms, both suitable for video feed from a web camera and images uploaded from storage devices. Thirdly, the design of an algorithm for calculation of the ratios between distances from iris corners was required, thus enabling real-time estimation of the angles. Lastly, a classifier based on Random Forest was developed, performing five types of eye detection: normal, esotropia, exotropia, hypertropia, and hypotropia, based on 16,654 images with balanced class weights. Importantly, the hybrid decision rule was designed in such a way that all obvious examples of squint eye syndrome would be recognized and less noticeable instances detected as well.

IV. METHODOLOGY

A. Dataset

Training of the model was done using the Strabismus dataset released by Ananthamoorthy on Kaggle [11]. The dataset contains 16,654 images of eye regions belonging to five categories. All images were loaded in grayscale format, scaled down to 60×30 pixels, and then flattened into a vector containing 1,800 elements. The distribution of classes can be found in Table I. Due to the higher number of normal instances compared to the strabismus subcategories, the Random Forest model was trained using class weights balanced.

TABLE I. DATASET CLASS DISTRIBUTION

Sr no.	Eye Condition	Number of Images	Percentage (%)
1.	Normal	10,154	60.97
2.	Esotropia	1,940	11.65
3.	Exotropia	1,635	9.82
4.	Hypertropia	1,530	9.19
5.	Hypotropia	1,395	8.38
	TOTAL	16,654	100.00

Pre-processing was done to standardize the image resolutions and facial alignment. In case of the imbalanced classes, balanced class weights were used in Random Forest.Where static images are involved, CLAHE technique was applied to the LAB color space of the images using a clip limit of 3.0 and 8x8 tiles. However, where live video images are involved, the images do not undergo this procedure in order to increase processing speed.

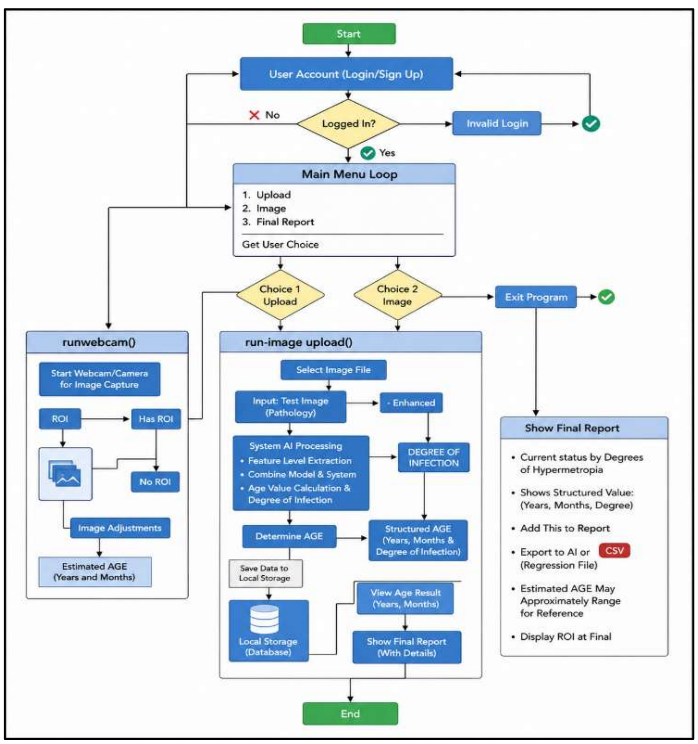


Fig. 1. System Flowchart

B. Landmark Detection and Feature Extraction

MediaPipe Face Mesh provides 478 facial landmarks in three dimensions per detected face, including more precise iris landmark points. The left iris center point is index 468 and the right iris center index is 473, while the left and right eye corners are located at indexes 133 and 33, 362 and 263 respectively. Eye region crops can be obtained by determining the bounding rectangle enclosing the four corner points, with an 8 pixel margin on all sides.The main geometric feature that is utilized is the iris-to-corner ratio, which is determined as the horizontal distance between the iris center and inner corner divided by the entire eye width. If the value is 0.5 then this implies the iris is centered, and is mapped into an angle using:

$$\text{Deviation (degrees)} = (\text{ratio} - 0.5) \times 100$$

The scaling factor of 100 has been selected experimentally to create substantial distinction between normal deviation and squint deviation readings. A negative output represents an inward movement of the iris, which can be indicative of esotropia, whereas a positive output represents an outward movement of the iris, indicating exotropia. If there is a substantial difference between both the eyes, then there is a vertical deviation, such as hypertropia or hypotropia.

C. Model Training

The model uses a Random Forest classifier with 200 trees having balanced class weights trained with 80 percent of the data set stratified into two portions. Stratification helps ensure that the training and testing sets represent each of the classes equally well and thus ensures that the testing set has a representation from all five classes. After being developed, the model can then be saved on the local disk using the pickle library in Python to make sure that it is not regenerated during application execution. Another classifier was also developed using Support Vector Machines.

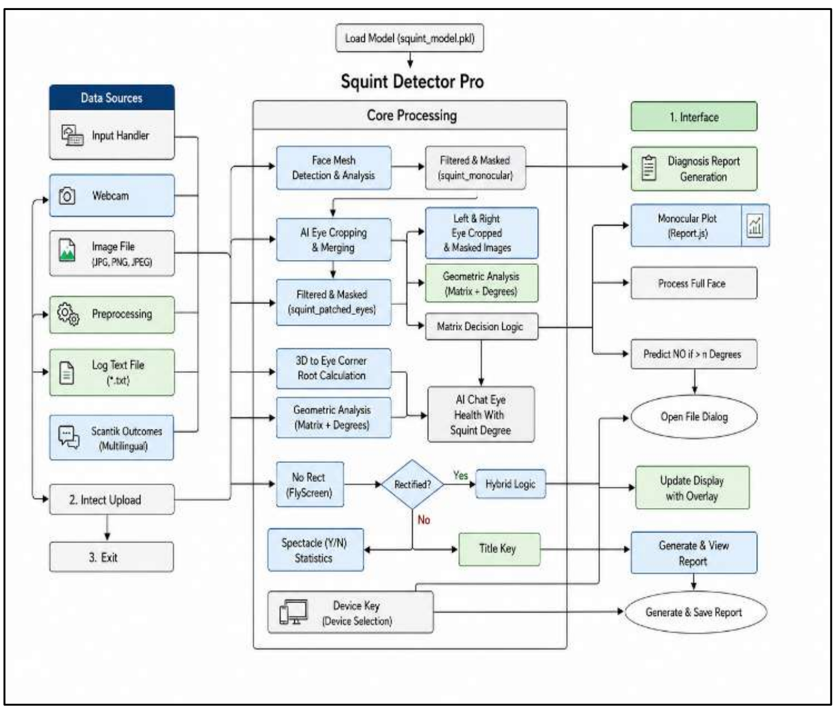


Fig. 2. System Architecture

D. Hybrid Decision Logic

Final classification is made on the basis of a two-tiered diagnostic approach. First, if the absolute value of the angle difference between the two eye deviations is greater than 7.0°, or an individual eye deviates more than 6.5°, then that particular case of strabismus is diagnosed based solely on geometry. Second, if any of the geometrical thresholds are not activated, then the output from the random forest algorithm will be checked, and a non-normal result having confidence more than 0.45 would constitute the final diagnosis.

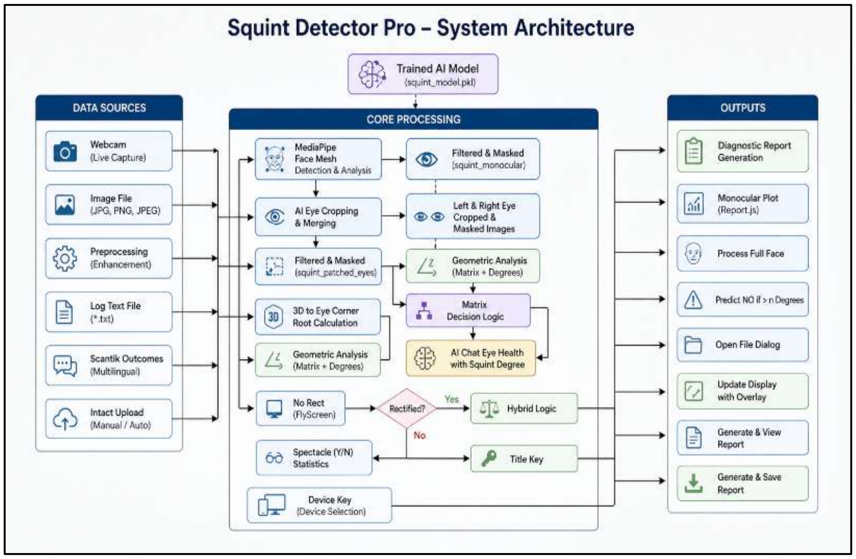


Fig. 3. Squint Detector Pro Architecture

E. Working Principle

Input into the model can be made either from the real-time webcam or an image uploaded through the web platform. In case of static image, CLAHE pre-processing helps highlight the irises regardless of different lighting conditions. Once the image is processed, it undergoes facial landmark analysis by MediaPipe Face Mesh where 478 3-D face landmarks are created. Iris centers and eye corner coordinates are used to calculate the ratio of iris to the corner point of the eyes. Ratio equals to 0.5 is indicative of perfectly straight alignment; if less than 0.5 then it is indicative of inward misalignment; and if greater than 0.5 then it is indicative of outward misalignment.

$$\text{Formula - Deviation} = (\text{ratio} - 0.5) \times 100.$$

While at it, the trimmed eye area is also classified using a pre-trained Random Forest model that gives out the predicted status alongside a corresponding confidence measure. The final decision about the patient's medical condition is made by employing a combination of two-stages logic. Where the geometrical anomaly surpasses specified values (6.5 degrees per eye and 7.0 degrees in disparity across eyes), strabismus is diagnosed immediately. But where the value from the classifier surpasses 0.45, the output from the classifier becomes the final decision. The results are presented together with deviation degrees and confidence value, while the program generates an automatic PDF diagnostic report.

V. RESULTS AND FINDINGS

The model was tested using the hold-out test set of 3,331 images, which constituted 20% of the entire data set. The Random Forest classifier performed well with an accuracy rate of 91.4% in classifying the test set images. The performance of SVM and Random Forest classifiers was compared by testing them using the same dataset. The Random Forest classifier performed better than the SVM classifier and was thus adopted in building the system, as shown in Table II.

TABLE II. MODEL PERFORMANCE COMPARISON

Model	Accuracy	F1 Score
Svm	85-88%	0.85
Random Forest	91.4%	0.91
Hybrid	Slight improvement	

This experiment has been run multiple times using stratified splitting, and similar results have been observed each time. The precision, recall, and F1 score for weighted precision, recall, and F1 score were 92%, 91%, and 91%, respectively. The average for all the classes together (macro average) was 83% for precision, recall, and F1 score.

Table III presents the complete per-class performance from the classification report obtained after training.

TABLE III. CLASSIFICATION REPORT — PER CLASS PERFORMANCE

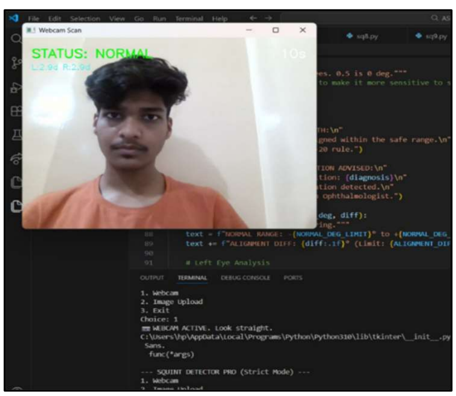
Sr No.	Condition	Precision	Recall	F1-Score	Support
1.	Esotropia	0.83	0.89	0.86	388
2.	Exotropia	0.73	0.79	0.76	327
3.	Hypertropia	0.77	0.79	0.78	306
4.	Hypotropia	0.85	0.67	0.75	279
5.	Normal	0.99	0.99	0.99	2,031
	Accuracy	-	-	0.91	3,331
	Macro Average	0.83	0.83	0.83	3,331
	Weighted average	0.92	0.91	0.91	3,331

The baseline class attained very high precision and recall at 0.99, proving that the model does not erroneously identify a normal eye as being squint. This high score is not as a result of class imbalance, as balanced class weights are incorporated. For the various squint classes, esotropia recorded the highest F1-score of 0.86 because the inward turning results in a unique iris texture in the eye region. Exotropia had an F1-score of 0.76, hypertropia 0.78, and hypotropia 0.75. The poor recall score of 0.67 for hypotropia implies that some observations were erroneously assigned to other classes.

Table IV presents the confusion matrix which gives a detailed breakdown of exactly where the misclassifications occurred across all five classes.

TABLE IV. CONFUSION MATRIX

Actual / Predicted	Esotropia	Exotropia	Hypertropia	Hypotropia	Normal
Esotropia	344	18	17	3	6
Exotropia	21	259	24	17	6
Hypertropia	14	39	241	12	0
Hypotropia	26	39	27	186	1
Normal	7	2	5	2	2,015



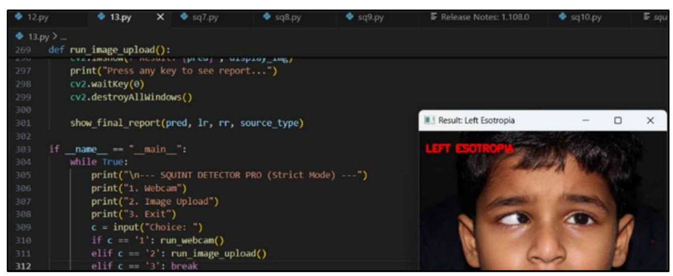


Fig. 4. Webcam Detection Output   Fig. 5. Esotropia Detection Result

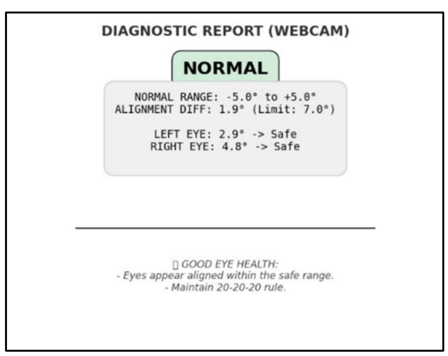


Fig. 6. Diagnostic Report Output

According to the confusion matrix, the most frequent errors made by the model include error in classifying adjacent strabismus types instead of error in classifying a strabismus type with respect to normal. Performance of the classification model is also analyzed via analysis using the classification report and confusion matrix. In some cases, exotropia is misdiagnosed as hypertropia and vice versa. The classification is clinically reasonable because a horizontal deviation could be interpreted as vertical deviation when seen from two-dimensional space like an image viewed in the front, while the cropped eye area (60x30 pixels) may not contain enough information for correct classification. Most importantly, very few squints are mistakenly classified as normal images. In the test set, 6 cases of esotropia, 6 cases of exotropia, 0 cases of hypertropia, and 1 case of hypotropia are classified falsely as normal images. The very low number indicates that very few real squints are mistaken as normal by the algorithm, which is very important for a diagnosis algorithm.
The results of geometric deviation analysis have yielded consistent values in the range of $-6.5°$ to $+6.5°$ for healthy eyes while cases of esotropia and exotropia were beyond this range. It was observed that the use of the hybrid logic model is more successful when either the classifier or the geometric analysis fails to detect a problem. The time taken for the whole operation is approximately 15 seconds per inference on a regular computer without a graphics processing unit. This makes the system applicable to brief screening sessions using one webcam capture only.

VI. DISCUSSION

The findings of the experiment showed that the suggested hybrid methodology was successful for detecting squint eyes. It can be seen that the performance of the Random Forest classifier was better than that of the Support Vector Machine, owing to the fact that the former could better account for the variability of the eye images. In terms of accuracy, it was noted that esotropia had the maximum F1-score, since it had unique visible features due to inward deviation of the iris; at the same time, hypotropia had the lowest value of recall due to small vertical deviations and lesser training data. In addition, the hybrid methodology made the decision process more reliable, combining the geometric deviation with machine learning classification. Moreover, the very few cases of squint being classified as non-squint indicate the future application of the system as a screening test.

VII. CONCLUSION

This research paper focused on the design, development, and implementation of an almost real-time system for squint eye detection based on geometric alignment of the irises along with Random Forest classification to form a hybrid decision model. Trained on a data set comprising of 16,654 images from five categories, our model obtained an accuracy rate of 91.4%, a weighted F1 score of 91%, and macro average F1 score of 83% from 3,331 test images. The confusion matrix reveals that there were almost no false negatives detected by the hybrid decision making model. The combination of the two processes means that the thresholds for geometrical deviations serve as the initial filter while the Random Forest deals with borderline cases. The developed algorithm takes webcam inputs as well as uploads images and provides PDF reports on the diagnostics. Further developments include the addition of a wider training set with clinically marked images, implementing head pose correction for increased stability, as well as creating a mobile application for rural areas. Nevertheless, clinical validation of the results is necessary for any recommendation within the healthcare sphere.

REFERENCES

[1] Kathpal, K., S. Negi, and S. Sharma. "iChat: Interactive Eyes for Specially Challenged People Using OpenCV Python." 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), Amity University, Noida, India (2021): 1-5.
[2] Kaur, J., and R. Ponnala. "Driver's Drowsiness Detection System Using Machine Learning." 2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT) (2023): 339-343.
[3] Mady, H., and S. M. S. Hilles. "Face Recognition and Detection Using Random Forest and Combination of LBP and HOG Features." International Conference on Smart Computing and Electronic Enterprise (ICSCEE) (2018): 1-7.
[4] Rohismadi, M. A. I., A. F. M. Raffei, N. S. A. Zulkifli, M. H. Ithnin, and S. F. Othman. "An Automated Strabismus Classification Using Machine Learning Algorithm for Binocular Vision Management System." 2023 IEEE 8th International Conference on Software Engineering and Computer Systems (ICSECS) (2023): 487-492.
[5] Gupta, B., M. Arora, S. Garg, and D. Ghosh. "Enhancing Strabismus Diagnosis from Detection to Classification with Deep Learning." 2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications (AIMLA) (2025): 1-6.

[6] Wang, J., and S. Jin. "Mobile Automated Strabismus Detection Using the Alternating Cover Test and Deep Learning." International Journal of Science and Research 14, no. 2 (2025): 231-238.
[7] Zhao, Z., H. Meng, S. Li, S. Wang, J. Wang, and S. Gao. "High-Accuracy Intermittent Strabismus Screening via Wearable Eye-Tracking and AI-Enhanced Ocular Feature Analysis." MDPI Biosensors 15, no. 2 (2025): 110.
[8] Ancona, C., R. Vera, M. Flores, and P. Gonzalez. "Methodology for Detecting Strabismus Through Video Analysis and U-NET Networks." IEEE Conference on Computer Systems (2024): 145-152.
[9] Lou, C., T. Zhang, W. Chen, and Y. Liu. "Comparative Analysis of Automated Strabismus Diagnosis Systems: End-to-End Versus Modular Architectures." IEEE Xplore International Conference Publication (2024): 89-96.
[10] Lu, Z., and M. Fan. "Deep Neural Network-Based Strabismus Diagnosis Through Smartphone Images for Telemedicine Eye Care." IEEE Xplore International Conference Publication (2024): 201-208.
[11] Ananthamoorthy, A. "Strabismus Dataset." Kaggle, 2023. https://www.kaggle.com/datasets/ananthamoorthya/strabismus.