Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Automated Detection and Classification of Strabismus in Smartphone Images

Anoop Thazhathumanackal Radhakrishnan
University of Waterloo
Brooke Koritala[2]
Retina Foundation of the Southwest
Robert F. Hess[3]
McGill University
Otman Basir[4]
University of Waterloo
Eileen E. Birch[2]
Retina Foundation of the Southwest
Benjamin Thompson[1]
ben.thomson@uwaterloo.ca
University of Waterloo

Research Article

Keywords: Artificial intelligence, Corneal light reflex, Deep learning, Mask R-CNN, Pseudo-strabismus, Smartphone imaging, Strabismus

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11437464/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Automated Detection and Classification of Strabismus in Smartphone Images

Original Research

Anoop Thazhathumanackal Radhakrishnan[1] · Brooke Koritala[2] · Robert F. Hess[3] · Otman Basir[4] · Eileen E. Birch[2,5*] · Benjamin Thompson[1,6*]

[1] School of Optometry and Vision Science, University of Waterloo, Waterloo, Ontario, Canada

[2] Retina Foundation of the Southwest, Dallas, Texas, USA

[3] McGill Vision Research, McGill University, Quebec, Canada

[4] Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, Canada

[5] Department of Ophthalmology, University of Texas Southwestern Medical Center, Dallas, Texas, USA

[6] School of Optometry and Vision Science, University of Waterloo, Waterloo, Ontario, Canada

[6]Centre for Eye and Vision Research, Hong Kong, China

*These authors contributed equally to this work.

Corresponding author: **Benjamin Thompson**
School of Optometry and Vision Science
University of Waterloo
200 University Avenue West
Waterloo, Canada.

Author emails: athazhat@uwaterloo.ca; bkoritala@rfsw.org; robert.hess@mcgill.ca; otman.basir@uwaterloo.ca; ebirch@retinafoundation.org; ben.thomson@uwaterloo.ca

ORCID iDs: Anoop Thazhathumanackal Radhakrishnan : 0009-0000-5603-8698
Robert F.Hess: 0000-0002-3324-2845
Otman Basir: 0000-0002-6454-0538
Eileen E. Birch: 0000-0001-8279-3291
Benjamin Thompson: 0000-0002-3902-2712

Abstract

Purpose: Accurate detection of strabismus from smartphone images remains challenging, particularly in distinguishing true strabismus from pseudo-strabismus, in which apparent ocular misalignment is caused by facial anatomy rather than abnormal eye position. This study developed and evaluated a fully automated smartphone-based approach that explicitly models interocular asymmetry between the corneal light reflex (CLR) and pupil center for strabismus detection and classification.

Methods: A Mask R-CNN framework was trained to localize pupil centers and CLR positions from facial images. Geometric asymmetry parameters derived from the CLR-pupil relationship in the two eyes were used to detect strabismus and classify detected cases as horizontal or vertical deviations, with horizontal cases further classified as esotropia or exotropia. The final evaluation used a separate set of 205 clinically labeled smartphone images comprising orthotropic children, pseudo-strabismus cases, and children with strabismus.

Results: Of 110 orthotropic images, 107 were correctly classified as orthotropic, while 87 of 95 strabismic images were correctly identified. Overall strabismus detection accuracy was 94.63%, with 96.67% precision, 91.58% sensitivity, 97.27% specificity, and an F1-score of 94.06%. Horizontal-versus-vertical classification achieved 95.40% accuracy, and classification of horizontal strabismus as esotropia or exotropia achieved 96.15% accuracy. None of the 12 pseudo-strabismus cases in the test set was misclassified as strabismus.

Conclusion: Explicit modeling of interocular CLR-pupil asymmetry enables accurate automated detection and classification of strabismus from smartphone images and may help distinguish true strabismus from pseudo-strabismus. Following prospective clinical validation, this approach has potential to support accessible smartphone-based strabismus screening.

Keywords: Artificial intelligence; Corneal light reflex; Deep learning; Mask R-CNN ; Pseudo-strabismus; Smartphone imaging; Strabismus

Key Summary Points

Why carry out this study?

- Accurate automated detection of strabismus from smartphone images remains challenging, particularly because image-based approaches may have difficulty distinguishing true strabismus from pseudo-strabismus.
- We investigated whether explicit modeling of the geometric relationship between the corneal light reflex and pupil center in both eyes could enable automated strabismus detection and subtype classification.

What was learned from the study?

- The proposed method achieved 94.63% accuracy, 91.58% sensitivity, and 97.27% specificity for strabismus detection, and all 12 pseudo-strabismus cases in the test set were classified as non-strabismic.
- The method classified horizontal versus vertical strabismus with 95.40% accuracy and horizontal deviations as esotropia or exotropia with 96.15% accuracy.
- Explicit CLR-pupil asymmetry provides an interpretable geometric basis for smartphone-based strabismus assessment and warrants prospective clinical validation for future screening use.

1. Introduction

Early detection of strabismus can enable timely treatment to reduce the risk of amblyopia and abnormal binocular function. In primary care and screening settings, ocular alignment is commonly assessed from the relative position of the corneal light reflex (CLR) between the two eyes using the Hirschberg test [1]. This assessment can be challenging in babies and young children outside specialized clinical eye-care settings. Smartphone imaging combined with automated image analysis may provide a practical approach to support strabismus assessment in clinical and community environments.

Current image-based methods for detecting strabismus use machine learning, deep learning, or spatial image-processing techniques [2-6]. Approaches based mainly on image appearance may be susceptible to misclassifying pseudo-strabismus as true strabismus because they do not explicitly represent the positional relationship between the pupil center and the CLR. In preliminary work, a VGG16-based convolutional neural network trained using images of non-strabismus (including pseudo-strabismus) and strabismus frequently misclassified pseudo-strabismus cases. This observation motivated the use of explicit CLR-pupil geometry as an interpretable feature for distinguishing true ocular misalignment from apparent misalignment caused by facial anatomy.

Several existing methods for strabismus detection also suffer from significant limitations in terms of automation, portability, and robustness to real-world imaging conditions. For instance, some methods [2] employ dedicated experimental setups that are unsuitable for mobile application development or fully automated workflows. Similarly, other methods [7–10] rely heavily on spatial image-processing techniques, often requiring specialized hardware or manual intervention, which may limit their scalability and practical deployment. For example, Hanan et al. [11] proposed a method based on video recordings taken before and after a cover test—an approach that inherently depends on manual testing procedures. Huang et al. [12] employed spatial techniques such as image binarization to detect light reflection points and iris regions; however, their method encounters difficulties with non-dark irises**,** low-resolution images, and subjects who wear spectacles. Moreover, it assumes the iris is perfectly centered within the palprebral fissure, which may not hold in many real-world scenarios.

Other researchers have explored distance-based detection strategies. Huang et al. [13] used the distance between the pupil center and facial landmarks to detect strabismus. However, the method is sensitive to variations in skin tone, illumination, facial morphology, and may fail to differentiate strabismus from pseudo-strabismus. Karaaslan et al. [14] focused on the distance between the corneal light reflex(CLR) and the pupil center. Although effective, their approach requires a predefined distance between the participant and a tripod-mounted camera, potentially limiting its use in busy clinical or screening environments.

Table 1 summarizes key differences between representative image-based strabismus detection methods and the proposed approach. Existing methods generally do not explicitly address pseudo-strabismus, provide limited strabismus subtype classification, or require controlled imaging conditions. In contrast, the proposed method combines smartphone-based image acquisition with explicit CLR–pupil asymmetry modeling to distinguish true strabismus from pseudo-strabismus while simultaneously classifying strabismus subtype.

Table 1 Comparison with Representative Image-Based Methods

Study	Deep learning	Explicit CLR-pupil modeling	Pseudo-strabismus detection	Subtype classification	Smartphone compatible
Chen et al. [2]	Yes	No	No	No	No
Huang et al. [13]	No	Partial	No	No	Yes
Karaaslan et al. [14]	Yes	Yes	No	No	Limited
Yarkheir et al. [5]	Yes	No	No	Partial	Yes
Proposed method	Yes	Yes	Yes	Yes	Yes

Our proposed method addresses these limitations through an approach that accurately identifies the CLR, the centre of the pupil and the relative distance between these points in images taken of both eyes by a smartphone under normal lighting conditions. The inter-ocular difference in distances enables detection of strabismus and classification of strabismus direction. The use of CLR and pupil features was intended to improve performance across variations in image acquisition conditions and ocular appearance while providing an interpretable geometric basis for strabismus assessment. Convolutional neural networks (CNNs) were employed to localize the CLR and pupil center accurately.

The primary contributions of this work are:

- Development of a smartphone-based framework that explicitly models interocular CLR–pupil asymmetry rather than relying solely on image-level classification.
- Introduction of geometric asymmetry parameters derived from CLR–pupil relationships to support discrimination between true strabismus and pseudo-strabismus.
- Automatic classification of detected strabismus into horizontal versus vertical deviations and further classification into esotropia and exotropia.
- A fully automated pipeline designed for single smartphone photographs acquired under normal indoor conditions without specialized ophthalmic imaging hardware or manual cover testing.

2. Methods

2.1. Proposal for Strabismus Detection

Figure 1 illustrates the proposed method. The process begins with face segmentation using the Haar cascade face detection algorithm[15]. The segmented face is then analyzed to estimate the pupil center position using a Mask R-CNN model trained on annotated face images. The CLR location is detected using the same trained Mask R-CNN model. The distance between the coordinates of the pupil center and CLR are calculated for both the left and right eyes. The algorithm determines the presence of strabismus if the difference in distance between the eyes exceeds a predefined threshold. Strabismus classification is achieved by analyzing the relative magnitude and sign of the pupil center and CLR distances in each eye.

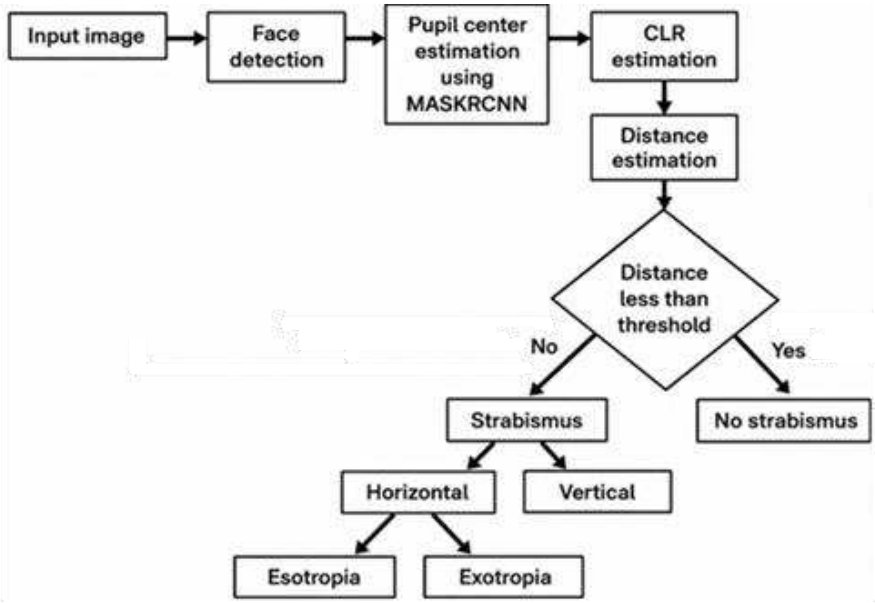


Fig.1. Block diagram illustrating the proposed method for strabismus detection and classification. CLR: corneal light reflection

2.2. Pupil Center and CLR Detection

Various methods have been developed to estimate pupil location in images or real-time video frames, particularly for gaze estimation. These methods often rely on machine learning algorithms or spatial image processing techniques to extract pupil coordinates. However, they tend to lack accuracy when dealing with low-resolution images, varying background conditions, or when spectacles are present. Additionally, in certain cases, only a small portion of the pupil may be visible in face images, making coordinate estimation more challenging. Traditional spatial image processing methods for pupil detection often fail under such circumstances. To address these limitations, the proposed method utilizes Mask R-CNN , a widely adopted instance segmentation framework introduced by He et al. [16], for accurate pupil center and CLR detection. Mask R-CNN was selected because accurate estimation of CLR and pupil centers requires pixel-level localization rather than object-level detection alone. Detection frameworks such as YOLO provide bounding boxes but do not directly generate segmentation masks. U-Net performs semantic segmentation but does not naturally separate multiple object instances. Media Pipe Iris is optimized for visible iris boundary detection and may fail when only partial iris visibility is present. Mask R-CNN provides instance-level segmentation with high localization accuracy, enabling robust estimation of CLR and pupil centroids under unconstrained imaging conditions. Three independent datasets were used for training, validation, and final testing, with complete subject-level separation among the datasets. The training dataset consisted of 200 facial images and was used to train the Mask R-CNN model. A separate validation dataset comprising 205 images was used to evaluate localization performance, optimize model parameters, and determine the geometric decision thresholds. Final performance was evaluated using a separate independent test dataset comprising 205 clinically labeled smartphone images. No image or participant was shared among the training, validation, and test datasets, thereby preventing data leakage between model development, validation, and final performance evaluation. The training dataset consisted of 200 facial images with a resolution of 1024 × 1024 pixels collected from publicly available online ophthalmology resources and image repositories. These images represented orthotropic subjects, pseudo-strabismus cases, and various forms of strabismus.

Pupil and corneal light reflex (CLR) regions were manually annotated using polygon masks to generate the ground-truth labels required for Mask R-CNN training. This annotation process involved marking the pixels corresponding to each pupil in the image. An example of an annotated image is shown in Figure 2, with both pupils highlighted with polygons. To further increase the diversity of the training set, data augmentation techniques were applied, including rotation, scaling, flipping, brightness adjustments, and the addition of noise to simulate real-world variations. The annotated data were organized into COCO JSON format, which is compatible with the Mask R-CNN training pipeline. The Mask R-CNN model was trained using this annotated dataset on Google Colab with GPU support, ensuring computational efficiency during training. The Mask R-CNN model employed a ResNet-50 backbone with Feature Pyramid Network (FPN) initialized using COCO pretrained weights. Training was performed using stochastic gradient descent with an initial learning rate of 0.001, momentum of 0.9, weight decay of 0.0001, and a batch size of two images. The model was trained for 100 epochs. Data augmentation included random rotations (±20°), horizontal flipping, scaling (0.8–1.2), brightness variation (±20%), and Gaussian noise. A confidence threshold of 0.5 was used during inference



Fig. 2 Example annotated image showing polygonal masks delineating the pupil regions

Optimization of the model's parameters minimized the difference between predicted and ground truth segmentation masks and bounding boxes. We evaluated the trained model on a separate validation set to assess its performance in segmenting pupils accurately. Fine-tuning was also performed by adjusting hyperparameters to improve performance. The centroid of the pixels in the detected pupil region gives the pupil coordinates. The centroid of a shape is the arithmetic mean (i.e. the average) of all the points in a shape. The image moment, which is defined as a set of statistical parameters to measure the distribution of pixel locations and their intensities, was used to identify the centroid of the segmented pupil region.

Mathematically the image moment Mij of order (i, j) for a grayscale image with pixel intensities I(x, y) is calculated as:

$$M_{i.j} = \sum_{x}\sum_{y} x^i y^j I(x.y) \qquad (1)$$

where x, y refer to row and column indexes and I(x, y) is the intensity of pixel at location (x, y).The centroid is given by the formula

$$Cx = M10/M00 \qquad (2)$$

$$Cy = M01/M00 \qquad (3)$$

Cx and Cy are the x and y coordinates of the centroid, M00 is the total number of non-zero pixels, M10 is the sum of all nonzero pixels(x coordinate) and M01 is the sum of all nonzero pixels(y coordinate). Figure 3 shows the pupil center denoted by red dots. The trained Mask R-CNN model detected the CLR and estimated its center for both the left and right eyes. The detection process employed techniques similar to those used for pupil center estimation.



Fig. 3 Detected pupil centers indicated by red dots

Figure 4 illustrates this process, showing two images: one displaying the bright spot on the eye (the CLR) and the other highlighting the detected CLR. The coordinates of the pupil center and CLR center were used to compute the distance between them. The distances in each eye were then compared to detect strabismus.



Fig. 4. Left: detected pupil centers and CLRS(bright white spots in each eye). Right: detected CLRs

The trained Mask R-CNN model was evaluated on a validation dataset consisting of 205 images obtained from kaggle strabismus image dataset. The pupil and CLR regions in the validation images were independently annotated using polygon masks to provide ground-truth segmentation labels for evaluation. The validation results are summarized in Table 2. Segmentation performance was assessed using Intersection over Union (IoU), precision, recall, and center localization accuracy. The model demonstrated strong performance for both pupil and corneal light reflex (CLR) detection. For pupil detection, the model achieved an IoU of 0.96, precision of 0.98, recall of 0.97, and correctly localized the pupil center in 200 of 205 validation images, corresponding to a center localization accuracy of 97.56%. For CLR detection, the model achieved an IoU of 0.94, precision of 0.96, recall of 0.95, and correctly localized the CLR center in 198 of 205 validation images, corresponding to a center localization accuracy of 96.59%. These results demonstrate that the proposed Mask R-CNN framework is capable of accurately localizing pupil and CLR regions under varying imaging conditions. The high localization performance enabled reliable extraction of CLR–pupil geometric features used for subsequent strabismus detection and classification. Although the localization accuracy was high, the final strabismus classification performance additionally depends on the geometric analysis of CLR–pupil asymmetry and threshold-based decision making, which may be affected by borderline deviations, image quality, and the presence of multiple corneal reflections

Table 2 Mask R-CNN localization performance on the validation set

Metric	Pupil Detection	CLR Detection
Validation Images	205	205
Correctly Localized Images	200/205	198/205
IoU	0.96	0.94
Precision	0.98	0.96
Recall	0.97	0.95
Center Localization Accuracy	97.56%	96.59%

Under normal lighting conditions there are often multiple corneal reflections present in addition to the camera/smartphone flash. The CLR caused by the flash can be characterized by its circular shape, distinguishing it from other non-circular regions such as room light reflections. To ensure detection of the correct CLR, we trained a Mask R-CNN model with images containing circular CLRs. This training enabled the model to effectively identify and distinguish the actual circular CLR from irrelevant reflections.We used circularity of the object detected to segment the correct CLR as follows. The trained Mask R-CNN was used to detect CLRs in new images. The output included segmented masks for all detected regions. For each detected region, circularity was calculated. Circularity can be defined as:

$$\textbf{Circularity} = \frac{\mathbf{4\pi \times Area}}{\mathbf{Perimeter^2}} \quad (4)$$

Here, the area is the number of pixels inside the detected region, and the perimeter is the number of pixels along the boundary of the region. The region with the highest circularity was selected as the CLR.

2.3. Estimation of Distance Between Pupil Centre and CLR

The face image was represented in a two-dimensional x-y coordinate system. The coordinates of the CLR in the right eye were (x_{clrR}, y_{clrR},0) and the left eye were(x_{clrL}, y_{clrL}, 0). The coordinates of the center of the pupil in the right eye were (x_{Rpc}, y_{Rpc},0) and in the left eye were(x_{Lpc}, y_{Lpc}, 0). The absolute horizontal and vertical displacements between the pupil center and CLR were calculated separately for each eye

The distance between the CLR and the center of the pupil for the left eye was

$$X_{difference_left} = abs(x_{Lpc} - x_{clrL}) \quad (5)$$

$$Y_{difference_left} = abs(y_{Lpc} - y_{clrL}) \quad (6)$$

The distance between the CLR and the center of the pupil for the right eye was

$$X_{difference_right} = abs(x_{Rpc} - x_{clrR}) \quad (7)$$

$$Y_{difference_right} = abs(y_{Rpc} - y_{clrR}) \quad (8)$$

Then the overall displacement in x direction was

$$D_x = abs(X_{difference_left} - X_{difference_right}) \quad (9)$$

Similarly overall displacement in y direction was

$$D_y = abs(Y_{difference_left} - Y_{difference_right}) \quad (10)$$

2.4. Strabismus Detection and classification

Strabismus detection was performed by comparing D_x and D_y against empirically determined thresholds. Threshold values for Dx and Dy were determined using Receiver Operating Characteristic (ROC) analysis performed on the validation dataset. Various threshold values were evaluated, and the corresponding sensitivity and specificity were computed. A threshold value of 3 pixels provided the optimal trade-off between sensitivity and specificity, yielding the

maximum Youden Index. Therefore, a threshold of 3 was selected for both Dx and Dy in the final strabismus detection algorithm. Subject-level separation was strictly maintained among the training, validation, and independent test datasets, with no participant appearing in more than one dataset. Consequently, threshold optimization was performed without subject-level data leakage. A value greater than 3 for either parameter indicated the presence of strabismus. If strabismus was detected, the deviation was classified as horizontal when $D_x > D_y + T_o$ and as vertical when $D_y > D_x + T_o$ where $T_o = 1$. For cases classified as horizontal, a second-stage classification was performed to determine whether the deviation was esotropia or exotropia. This classification utilized the signed horizontal displacement between the pupil center and the CLR. For each eye, the horizontal displacement was computed as:

$$d_x^{(L)} = x_{Lpc} - x_{clrL} \quad (11)$$

$$d_x^{(R)} = x_{Rpc} - x_{clrR} \quad (12)$$

The eye with the greater absolute horizontal displacement, |dx(L)| or |dx(R)|, was designated as the deviating eye. The sign of the horizontal displacement of the dominant eye was used to determine the horizontal strabismus subtype. In the frontal facial images used in this study, the eye appearing on the left side of the image corresponds to the participant's right eye, whereas the eye appearing on the right side corresponds to the participant's left eye. The horizontal displacement was computed using Eqs. (11) and (12) in the standard image coordinate system, where the x-axis increases from left to right. Because the anatomical temporal and nasal directions correspond to opposite image-coordinate directions for the two eyes, the sign of the horizontal displacement was interpreted according to eye laterality. Specifically, for the participant's right eye (left side of the image), $dx < 0$ indicates temporal displacement and $dx > 0$ indicates nasal displacement. Conversely, for the participant's left eye (right side of the image), $dx > 0$ indicates temporal displacement and $dx < 0$ indicates nasal displacement**.** The anatomical direction of the dominant horizontal displacement was then used to classify the strabismus subtype. A temporal displacement of the dominant eye was classified as exotropia (outward deviation), whereas a nasal displacement was classified as esotropia (inward deviation).

3. Results

3.1. Strabismus Detection

The testing dataset consisted of 205 clinically labeled facial images collected using a handheld smartphone at the Pediatric Vision Laboratory of the Retina Foundation of the Southwest, Dallas, TX, USA. The testing dataset included orthotropic subjects, pseudo-strabismus cases, and clinically diagnosed strabismus cases. All clinical labels were assigned by experienced pediatric ophthalmologists. Images were acquired using handheld smartphones under normal indoor illumination with the integrated camera flash enabled. The camera-to-subject distance was adjusted so that the participant's face was fully visible within the image frame. Images were resized to 1024 × 1024 pixels before processing. Detection results are shown in Table 3. All pseudo-strabismus cases were correctly classified as non-strabismic. The proposed method achieved an overall accuracy of 94.63% with 96.67% precision for detecting strabismus. Recall (sensitivity) was 91.58% and specificity was 97.27%. The combined performance measured using the F1-score was 94.06%, confirming the balanced performance of the proposed approach in terms of both precision and recall.

Table 3. Detection Results

Actual/Predicted	**Orthotropic**	**Strabismic**
Orthotropic(110)	107(97.3%)	3(2.7%)
Strabismic (95)	8(8.4%)	87(91.6%)

To quantify statistical uncertainty, 95% confidence intervals were estimated using bootstrap resampling with 1000 iterations. The overall accuracy was 94.63% (95% CI: 91.22%–97.56%), sensitivity was 91.58% (95% CI: 85.57%–96.74%), and specificity was 97.27% (95% CI: 93.91%–100.00%). The corresponding 95% bootstrap confidence intervals are presented in Table 4

Table 4. Strabismus detection performance with 95% bootstrap confidence intervals

Metric	**Value in %**	**95% Bootstrap CI (%)**
Accuracy	94.63	91.22 – 97.56
Sensitivity	91.58	85.57 – 96.74
Specificity	97.27	93.91 – 100.00

3.2. Strabismus Classification

The proposed algorithm was further evaluated for the classification of strabismus into horizontal and vertical subtypes. Only correctly classified strabismus images were used (n = 87, Table 3). The confusion matrix is shown in table 5.

Table 5. Horizontal and vertical strabismus classification results

Actual/Predicted	**Horizontal**	**Vertical**
Horizontal(82)	78(95.1%)	4(4.9%)
Vertical(5)	0(0.0%)	5(100%)

The proposed model achieved an overall accuracy of 95.40%. Treating horizontal strabismus as the positive class, sensitivity was 95.1% and specificity was 100%. And F1 score was 97.50%. Incorrect classifications were due to low resolution (photograph taken from a long distance), lack of a CLR in one eye, too many potential CLR points and glare due to spectacles.

3.3. Classification of Horizontal Strabismus into esotropia and exotropia

This analysis was conducted using the 78 horizontal strabismus images correctly classified in the preceding stage (Table 5). Treating esotropia as positive class, classification results are shown in Table 6. The method achieved an overall accuracy of 96.15%, with a recall of 97.73%, specificity of 94.12% and an F1 score 96.61%.

Table 6. Esotropia and exotropia classification results

Actual/Predicted	**Esotropia**	**Exotropia**
Esotropia(44)	43(97.7%)	1(2.3%)
Exotropia(34)	2(5.9%)	32(94.1%)

4. Discussion

4.1 Principal Findings

This study developed and evaluated a fully automated smartphone-based approach for strabismus detection and classification that combines deep-learning-based localization of the pupil and corneal light reflex (CLR) with explicit geometric analysis of interocular CLR–pupil asymmetry. On the independent clinical test set of 205 smartphone images, the proposed method achieved an overall strabismus detection accuracy of 94.63%, sensitivity of 91.58%, specificity of 97.27%, precision of 96.67%, and an F1-score of 94.06%. Bootstrap analysis further demonstrated relatively narrow confidence intervals for the principal detection metrics, with an accuracy of 94.63% (95% CI 91.22–97.56%), sensitivity of 91.58% (95% CI 85.57–96.74%), and specificity of 97.27% (95% CI 93.91–100.00%).

Beyond binary strabismus detection, the proposed geometric approach enabled classification of the direction and subtype of ocular deviation. Among the 87 correctly detected strabismic images, horizontal-versus-vertical classification achieved an accuracy of 95.40%. Among the 78 correctly classified horizontal cases, subsequent classification as esotropia or exotropia achieved an accuracy of 96.15%. These findings indicate that the relative position of the pupil center and CLR contains useful information not only for detecting ocular misalignment but also for characterizing the direction of deviation.

An important observation was that none of the 12 pseudo-strabismus cases included in the independent test set was classified as strabismic. Pseudo-strabismus presents a particular challenge for image-based approaches because facial anatomical characteristics can create the appearance of ocular misalignment despite normal ocular alignment. Preliminary observations using a VGG16 image-level classification approach motivated the present geometric strategy because pseudo-strabismus cases could be misclassified when decisions were based predominantly on overall image appearance. By explicitly localizing the pupil center and CLR in each eye and comparing their relative positions, the proposed method bases its decision on features more directly related to ocular alignment. Although the number of pseudo-strabismus cases was limited, the absence of false-positive classifications in this subgroup provides encouraging preliminary evidence supporting this design rationale.

4.2 Comparison with Existing Approaches

The proposed method differs from several previously reported image-based approaches [2,5,13,14] by integrating smartphone-compatible image acquisition, deep-learning-based landmark localization, explicit CLR–pupil geometric analysis, pseudo-strabismus discrimination, and strabismus subtype classification within a single automated framework.

Previous studies have investigated CNN-based approaches for strabismus recognition [2,5]. Although deep-learning classifiers can learn discriminative image features directly from facial or ocular images, image-level classification does not necessarily provide an explicit representation of the anatomical relationship between the pupil and CLR that underlies conventional assessment of ocular alignment. This distinction may be particularly relevant in pseudo-strabismus, where facial morphology can produce the visual appearance of misalignment despite normal ocular position.

Other approaches have used spatial or distance-based measurements. Huang et al. [13], for example, used relationships between the pupil center and facial landmarks. Such measurements may be affected by variations in facial morphology, illumination, and image acquisition conditions and do not directly represent the relationship between the CLR and pupil center. Karaaslan et al. [14] investigated CLR–pupil distance measurements but employed more controlled camera positioning. In contrast, the present method was designed for photographs acquired using a handheld smartphone under normal indoor illumination, thereby reducing dependence on specialized imaging equipment or fixed camera positioning.

Another distinction of the proposed framework is that it extends beyond binary detection. Following identification of strabismus, interocular horizontal and vertical asymmetry is used to distinguish horizontal from vertical deviations, and signed horizontal CLR–pupil displacement is subsequently used to differentiate esotropia from exotropia. The resulting pipeline therefore provides an interpretable progression from landmark localization to detection and subsequent characterization of the deviation.

4.3 Clinical and Technical Implications

The combination of deep-learning-based localization and explicit geometric measurements provides an interpretable framework for automated strabismus assessment. Rather than producing only an image-level class prediction, the method identifies the pupil center and CLR and derives the classification from their spatial relationship between the two eyes. This provides a direct connection between the computational decision process and ocular alignment features that are relevant to established clinical assessment methods such as evaluation of the corneal light reflex.

Smartphone-compatible image acquisition may also be advantageous for future screening applications. Smartphones are widely available and offer a potentially accessible platform for image acquisition outside specialized ophthalmic imaging environments. A system capable of automatically analyzing a single facial photograph could eventually complement existing screening approaches in community, primary-care, or remote settings.

However, the results of the present study should be interpreted as an evaluation of a research method rather than evidence of readiness for autonomous clinical diagnosis. The algorithm was evaluated retrospectively on a defined clinical image dataset. Before clinical deployment, prospective validation is required in larger and more diverse populations and under the range of acquisition conditions expected in real-world screening. The potential role of the method would also need to be evaluated within an appropriate clinical workflow, including procedures for referral and professional assessment following a positive screening result.

4.4 Strengths and Limitations

Several aspects of the study strengthen the evaluation of the proposed approach. Separate training, validation, and independent testing datasets were used, with subject-level separation to minimize the risk of data leakage. The final evaluation was conducted using 205 clinically labelled smartphone images, with clinical labels assigned by experienced pediatric ophthalmologists. The test set included orthotropic, strabismic, and pseudo-strabismus cases, enabling

assessment of a clinically relevant source of potential false-positive classifications. Statistical uncertainty for the principal detection metrics was also evaluated using bootstrap confidence intervals. In addition, performance was assessed beyond binary detection through classification of horizontal and vertical deviations and subsequent differentiation of esotropia and exotropia.

The study also has several limitations. First, although the independent test set contained 205 images, larger multicentre datasets will be required to establish generalizability across different populations, devices, imaging environments, and clinical settings. In particular, only 12 pseudo-strabismus cases were included in the test set. The correct classification of all 12 cases is encouraging but should therefore be considered preliminary evidence rather than definitive validation of pseudo-strabismus discrimination.

Second, only five vertical-strabismus cases were included among the 87 correctly detected strabismic images used for directional classification. Although all five were correctly classified as vertical deviations, the small sample size prevents strong conclusions regarding performance for vertical strabismus. Further evaluation using substantially larger numbers and a broader range of vertical deviations is required.

Third, image quality influenced algorithm performance. Incorrect classifications were associated with low-resolution images, absence of a clearly detectable CLR in one eye, multiple potential CLR points, and reflections or glare associated with spectacles. These factors are particularly relevant to smartphone-based acquisition because image quality, subject positioning, ambient illumination, camera characteristics, and flash reflections may vary considerably outside controlled research settings.

Fourth, the proposed decision process depends on empirically determined geometric thresholds. Although these thresholds were selected using the separate validation dataset and subsequently applied to the independent test set, additional external validation is required to determine whether the same thresholds generalize across different populations, smartphone models, image resolutions, and acquisition conditions.

Finally, the current study evaluates strabismus from individual frontal facial photographs and does not capture the full complexity of clinical ocular alignment assessment. Strabismus can vary with gaze direction, viewing distance, fixation conditions, and intermittency. The present approach should therefore be considered a potential screening methodology rather than a replacement for comprehensive examination by an eye-care professional.

4.5 Future Work

Future work will focus on prospective evaluation using larger and more diverse clinical populations, with particular emphasis on increasing the number of pseudo-strabismus and vertical-strabismus cases. External validation across different clinical centres, smartphone models, age groups, and imaging environments will be important for determining the generalizability of the proposed approach.

Automated image-quality assessment could also be incorporated into the mobile workflow. Before strabismus analysis, the system could assess factors such as face position, image resolution, visibility of both pupils and CLRs, excessive

glare, and the presence of multiple reflections. When an image does not satisfy predefined quality requirements, the user could be prompted to reacquire the photograph rather than generating a potentially unreliable result.

Further development could also evaluate real-time implementation within a smartphone application and investigate performance across multiple gaze positions. Prospective comparison with established clinical ocular alignment assessments will ultimately be required to determine the sensitivity, specificity, usability, and potential clinical utility of the approach in real-world screening workflows.

5. Conclusion

This study presents an automated smartphone-based approach for detecting and classifying strabismus using deep-learning-based localization of the pupil and corneal light reflex combined with explicit modelling of interocular CLR–pupil asymmetry. On an independent clinical test set of 205 smartphone images, the method achieved 94.63% accuracy, 91.58% sensitivity, and 97.27% specificity for strabismus detection. The framework also demonstrated high accuracy for horizontal-versus-vertical classification and subsequent differentiation of esotropia and exotropia.

The correct classification of all 12 pseudo-strabismus cases as non-strabismic provides encouraging preliminary evidence that explicit CLR–pupil geometry may help distinguish true ocular misalignment from apparent misalignment associated with facial anatomy. However, larger prospective and external validation studies, particularly those including greater numbers of pseudo-strabismus and vertical-strabismus cases, are required. Following such validation, the proposed approach may provide a foundation for accessible smartphone-based strabismus screening and referral support.

Acknowledgment

This work was supported by a CIHR project grant to BT, EB, RFH and OB. NSERC grants to BT, and RFH.

Medical Writing, Editorial, and Other Assistance

No medical writing or editorial assistance was declared in the source manuscript

Authorship

All named authors should meet the journal's authorship criteria and approve the final manuscript.

Anoop Thazhathumanackal Radhakrishnan, Brooke Koritala, Robert F. Hess, Otman Basir, Eileen E. Birch,and Ben Thompson contributed equally to the conception and design of the study, development of the methodology, analysis and interpretation of the results, and preparation and critical revision of the manuscript. All authors reviewed and approved the final version of the manuscript and agree to be accountable for the work.

Conflict of Interest

Authors have no conflicts of interest if that is accurate

Ethics Approval

The study protocol was reviewed and approved by the University of Waterloo Research Ethics Board. All study procedures were conducted in accordance with the applicable institutional ethical standards and the principles of the Declaration of Helsinki.

Competing Interests

The authors declare that they have no competing financial or non-financial interests directly or indirectly related to the work submitted for publication

References

1. Loh AR, Chiang MF. Pediatric vision screening. Pediatr Rev. 2018;39(5):225-234. doi:10.1542/pir.2016-0191.
2. Chen Z, Fu H, Lo WL, Chi Z. Strabismus recognition using eye-tracking data and convolutional neural networks. J Healthc Eng. 2018;2018:1-9. doi:10.1155/2018/7692198.
3. De Figueiredo LA, Dias JVP, Polati M, Carricondo PC, Debert I. Strabismus and artificial intelligence app: optimizing diagnostic and accuracy. Transl Vis Sci Technol. 2021;10(7):22. doi:10.1167/tvst.10.7.22.
4. Mao K, et al. An artificial intelligence platform for the diagnosis and surgical planning of strabismus using corneal light-reflection photos. Ann Transl Med. 2021;9(5):374. doi:10.21037/atm-20-5442.
5. Yarkheir M, Sadeghi M, Azarnoush H, Akbari MR, Khalili Pour E. Automated strabismus detection and classification using deep learning analysis of facial images. Sci Rep. 2025;15(1):3910. doi:10.1038/s41598-025-88154-6.
6. Zhang G, et al. Multi-feature fusion-based strabismus detection for children. IET Image Process. 2023;17(5):1590-1602. doi:10.1049/ipr2.12740.
7. Model D, Eizenman M, Sturm V. Fixation-free assessment of the Hirschberg ratio. Invest Ophthalmol Vis Sci. 2010;51(8):4035. doi:10.1167/iovs.09-5014.
8. Lim H.W., et al. Clinical measurement of the angle of ocular movements in the nine cardinal positions of gaze. Ophthalmology. 2014;121(4):870-876. doi:10.1016/j.ophtha.2013.11.019.
9. Barry JC, Effert R, Reim M, Meyer-Ebrecht D. Computational principles in Purkinje I and IV reflection pattern evaluation for the assessment of ocular alignment. Invest Ophthalmol Vis Sci. 1994;35(13):4205-4218.
10. Yang HK, Seo JM, Hwang JM, Kim KG. Automated analysis of binocular alignment using an infrared camera and selective wavelength filter. Invest Ophthalmol Vis Sci. 2013;54(4):2733. doi:10.1167/iovs.12-11400.
11. Abdullah Mengash H, Hosni Mahmoud HA. Methodology for detecting strabismus through video analysis and intelligent mining techniques. Comput Mater Contin. 2021;67(1):1013-1032. doi:10.32604/cmc.2021.014942.
12. Huang X-L, Zoo K, Han C. An automatic strabismus screening method with corneal light reflex based on image processing. J Korea Multimed Soc. 2021;24(5):642-650.
13. Huang X, Lee SJ, Kim CZ, Choi SH. An automatic screening method for strabismus detection based on image processing. PLoS One. 2021;16(8):e0255643. doi:10.1371/journal.pone.0255643.
14. Karaaslan Ş, Kobat SG, Gedikpınar M. A new method based on deep learning and image processing for detection of strabismus with the Hirschberg test. Photodiagnosis Photodyn Ther. 2023;44:103805. doi:10.1016/j.pdpdt.2023.103805.
15. Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). Kauai, HI, USA: IEEE Computer Society; 2001. p. I-511-I-518. doi:10.1109/CVPR.2001.990517.
16. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN . arXiv. 2018;arXiv:1703.06870. doi:10.48550/arXiv.1703.06870.