Original Article

Effectiveness of Kanna photoscreener in detecting amblyopia risk factors

Kaushik Murali, Viswesh Krishna[1], Vrishab Krishna[2], B Kumari, Sowmya Raveendra Murthy, Vidhya C, Payal Shah

Purpose: Amblyopia is a significant public health problem. Photoscreeners have been shown to have significant potential for screening; however, most are limited by cost and display low accuracy. The purpose of this study was validate a novel artificial intelligence (AI) and machine learning–based facial photoscreener "Kanna," and to determine its effectiveness in detecting amblyopia risk factors. **Methods:** A prospective study that included 654 patients aged below 18 years was conducted in our outpatient clinic. Using an android smartphone, three images of each the participants' face were captured by trained optometrists in dark and ambient light conditions and uploaded onto Kanna. Deep learning was used to create an amblyopia risk score based on our previous study. The algorithm generates a risk dashboard consisting of six values: five normalized risk scores for ptosis, strabismus, hyperopia, myopia and media opacities; and one binary value denoting if a child is "at-risk" or "not at-risk." The presence of amblyopia risk factors (ARF) as determined on the ophthalmic examination was compared with the Kanna photoscreener. **Results:** Correlated patient data for 654 participants were analyzed. The mean age of the study population was 7.87 years. The algorithm had an F-score, 85.9%; accuracy, 90.8%; sensitivity, 83.6%; specificity, 94.5%; positive predictive value, 88.4%; and negative predictive value, 91.9% in identifying amblyopia risk factors. The P value for the amblyopia risk calculation was 8.5×10^{-142} implying strong statistical significance. **Conclusion:** The Kanna photo-based screener that uses deep learning to analyze photographs is an effective alternative for screening children for amblyopia risk factors.

Key words: Amblyopia, deep learning, mobile phone, screening

Amblyopia is the leading cause of childhood visual impairment in the world, affecting 1% to 6% of all children, with nearly 100 million children suffering from it.[1,2] If left untreated, amblyopia could result in complete vision loss in one or both eyes.[3] Furthermore, amblyopia treatment is limited by age (visual maturation period), and so early detection is crucial for a successful recovery.[4] Thus, vision screening is recommended for children less than 3 years of age in the United States and less than 5 years in India.[5,6] Traditional vision screening typically involves visual acuity testing and refractive error measurement, both of which can be quite challenging and time-consuming due to the poor cooperation of children in this age-group during such examinations.[1] Moreover, traditional methods are felt to have high overreferral rates, low sensitivity, and low specificity.[6] Thus, there is a need for more effective screening methodologies.

Although the AAP (American Academy of Pediatricians) recommends photoscreening in children under the age of 5 years to positively identify amblyopia risk factors, this is practically difficult to achieve in India due to the limited resources and infrastructure.[7,8] Thus, most at-risk children, especially in resource-starved regions are not screened in time.

Instrument-based infrared ARF photoscreening systems such as Plusoptix (Plusoptix GmbH, Germany), 2WIN (Adaptica, Italy), and SPOT (Welch Allyn, USA) have been shown to identify the risk factors that are likely to lead to amblyopia or poor vision.[9-11] They hold several advantages as compared with the traditional methods: They are faster, more objective, and require less cooperation from children. However, the high acquisition and management costs associated with such instruments have prevented their widescale use as a screening tool.[12]

Smartphone-based ARF screening systems have been shown to be useful in predicting specific ocular disorders, such as diabetic retinopathy and leukocoria, and in vision screening.[13,14] GoCheckKids (Gobiquity Inc., USA) is a smartphone-based vision screening system for children that uses an iPhone 7 or Nokia Lumia and a customized phone case with built-in image processing software that identifies certain amblyopia risk factors.[15] This system has been validated clinically in several studies[14,16-18] that analyze two separate detection methods: an automated algorithm and manual grading of photographs acquired from the smartphone system. Manual grading of photographs for GoCheckKids increases the sensitivity of the detection methodology significantly.[17]

Sankara Academy of Vision, Sankara Eye Hospital, [1]Stanford University, Stanford, California, USA, [2]National Public School Bengaluru, Karnataka, India

Correspondence to: Dr. Kaushik Murali, Sankara Eye Hospital, Varthur Road, Kundalahalli Gate, Bangalore - 560 037, Karnataka, India. E-mail: kaushik@sankaraeye.com

Our previous work described "Kanna," a deep learning and image processing approach for detection of amblyopia risk factors using smartphones.[19]

Kanna is based on the acquisition of two images to predict the presence of ARF: one in normal ambient light conditions to help predict ptosis and strabismus and a low light image of the face to help identify media opacities, anisometropia, and isometropia.

To reduce variation from acquisition, the eye was localized using facial landmarks predicted by DL models. A convolutional neural network (CNN) was trained to detect six iris landmarks along the iris boundary using the UnityEyes data set [Fig. 1].[19,20] An android application was written to modify the smartphone camera flash settings to remove the inbuilt preflash that removes red-eye images in low-light conditions.

To predict refractive ARF, Kanna uses the methods of eccentric photorefraction [Fig. 2].[19,21] Using the width of the bright crescent that occurs on one side of the reflex, Kanna calculates the refractive error of the eye. Depending on the side of the pupil at which the crescent is located, it can be classified as hyperopia or myopia (see Bobier and Braddick for a more detailed treatment).[21]

We had shown that such a methodology held significant promise as a screening system based on a pilot study with 50 individuals in Murali *et al*. This work is a clinical validation and analysis of our previously proposed system.[19]

Methods

This was a prospective study that included 654 patients aged below 18 years, conducted in the pediatric ophthalmology outpatient department of Sankara Eye Hospital, Bangalore, India, from December 2018 to September 2019, after approval from the scientific and ethics review board of the institution. Oral and written consent was obtained from the parent/ guardian of the patient before recruiting in the study.

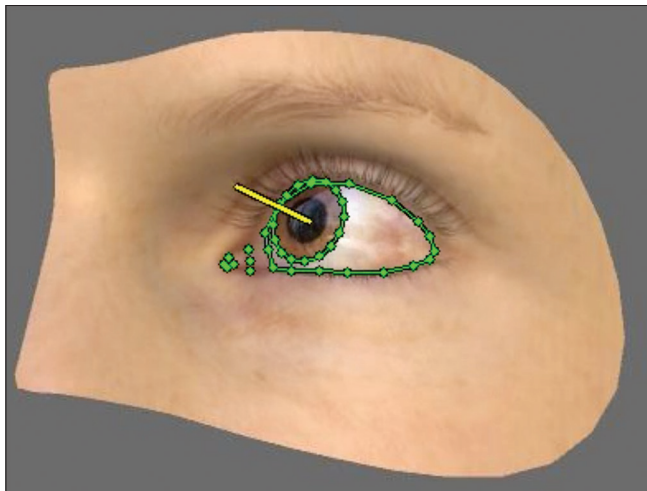The sample size was calculated with a 5% level of significance, 90% power, 6% prevalence of amblyopia recorded in the outpatient department, and a 3% maximum error limit of detection of amblyopia using a hand-held device.

$$N = \frac{(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2 \times P \times (1-P)}{e^2}$$

α = 5% level of significance $\Rightarrow Z_{1-\frac{\alpha}{2}} = 1.96$

β = 90% power of study $\Rightarrow Z_{1-\beta}) = 1.28$

P = 6% prevalence of amblyopia

e = 3 % error limit of detection

Data acquisition

An Android-based smartphone (OnePlus 6T) was used for collecting images of the patients. A specialized android application described in Murali *et al.* was used on this phone for reliable and reproducible generation of red reflex and ambient images.[19] This screening system was compared with a comprehensive eye exam by one of the coauthors (KM, SR, VC) as the gold standard.

As part of the smartphone-based screening, three images of the participant's face were captured by trained optometrists in a closed room. The first two images were captured in darkened conditions (3–10 lm) at a distance of approximately 1 m. Of these two images, the first image was captured holding the smartphone horizontally with the camera closer to the left hand of the screener, and the second image was captured with the smartphone held vertically. The third image was captured in normal ambient light conditions (60–800 lm) at a distance of approximately 0.5 m. These requirements were arrived at based on testing different light conditions and distances as described in Murali *et al.*[19] This served to reduce interobserver variability by ensuring standard operating procedures for image capture.

Next, a comprehensive eye examination was conducted by a pediatric ophthalmologist and an optometrist. This included cover tests for distant and near ocular movements, detailed anterior segment examination, cycloplegic refraction, and fundus examination in the same order as mentioned. Data on the presence of any amblyopia risk factor such as ptosis, strabismus, refractive error, or media opacity were noted.
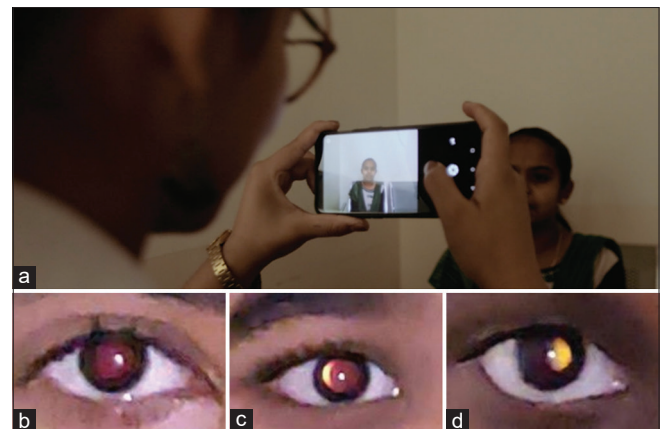

Figure 1: Generated eye image with iris contour and landmarks (from https://bit.ly/2OjnbIH)

Figure 2: Red reflex images: (a) Low light image acquisition (b) Emmetropic red reflex (c) Hyperopic red reflex (d) Myopic red reflex

A unique photograph identification number was used to correlate the patient examination with the algorithm data.

Data processing

In practice, vision screening systems are required to predict if a participant is "at-risk" and should be referred to an ophthalmologist for further testing or if the child is "not at-risk." Our algorithm generates a risk dashboard based on the 2003 AAPOS (American Association for Pediatric Ophthalmology and Strabismus) guidelines[22] consisting of six values: five normalized risk scores for ptosis, strabismus, hyperopia, myopia, and media opacities, and one binary value denoting if a child is at-risk or not at-risk. The patient data were processed according to the 2003 AAPOS guidelines to compare effectively with the risk dashboard. The 2003 AAPOS referral criteria are provided for the readers' convenience:

- Anisometropia (spherical or cylindrical) >1.5D
- Hyperopia >3.5D in any meridian
- Myopia >3.0D in any meridian
- Astigmatism >1.5D at 90° or 180°; >1.0D in oblique axis (more than 10° from 90° or 180°)
- Any manifest strabismus
- Any media opacity >1 mm in size
- Ptosis ≤1 mm margin reflex distance.

The images were anonymized by using facial landmark detection models to extract the eyes of the participant and were uploaded to a secure cloud storage location for processing. Each participant's ambient and horizontal low-light photographs were run through the detection algorithm. In cases where the algorithm was not able to process one or more of the images, an "at-risk" prediction was assigned to indicate that further testing was required. This approach was used to accommodate cases with large ptosis or strabismus where red reflex images could not be processed.

Microsoft Excel and the NumPy, SciPy, and Pandas libraries in the Python programming language were used for statistical analysis.[23-25] Phi Coefficients were calculated for individual risk factors and overall amblyopia risk prediction. Based on the degree of freedom (DF) and sample size, a *P* value at a 0.05 significance threshold was calculated. Several other metrics such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F-score were calculated based on the confusion matrix.

Results

A total of 802 participants were screened, of which 148 either revoked consent during or after the process or did not cooperate, leaving 654 participants with correlated patient data.

The participants were between the ages of 11 months to 18 years with a mean age of 7.87. The male–to-female ratio was 1.17:1. The number of participants and phi coefficients for each amblyopia risk factor can be found in Table 1. The imageability rate (percentage of respondents where smartphone-based screening succeeded) was 97.4%. We find that amblyopia and strabismus predictions are strongly correlated and the other risk factors are moderately correlated.

The confusion matrix for amblyopia risk calculation can be found in Table 2. The *P* value for the amblyopia risk calculation

was 8.5×10^{-142}, which is much smaller than 0.05, meaning the results are strongly statistically significant.

The sensitivity and specificity of identifying amblyopia risk factors by our algorithm (Kanna) are 83.6% and 94.5%, respectively. Out of 654, there were 24 false positives that were either on the verge of being amblyogenic or had a slight error in calculation and 36 false negatives that were mainly due to the inability to get a clear red reflex in these cases. The F-score value was 85.9% with a positive predictive value of 88.4% and a negative predictive value of 91.9%. The accuracy metrics with 95% confidence intervals can be found in Table 3.

Discussion

The proposed vision screening system has high accuracy for the identification of ARF and detects individual risk factors with moderate to high accuracy using a smartphone application.

Although the number of cases with ptosis and media opacity risk was limited, they still have a moderate correlation with clinical prediction. The medium correlation of anisometropia and isometropia is partially due to cases with excessive ptosis or strabismus where red reflex analysis is skipped and set to "no risk" by default. Media opacity detection could be increased by utilizing specialized models for detecting different media opacities.

Table 1: Data set and prediction composition with phi coefficients

	Clinical	Kanna	Phi Coefficient
Amblyopia	219	207	0.79
Ptosis	11	3	0.42
Strabismus	56	56	0.79
Anisometropia	71	87	0.43
Isometropia	129	61	0.51
Media Opacities	16	17	0.47

Table 2: Confusion matrix for clinical validation

	Kanna Positive	Kanna Negative	Row Total
Clinical Positive	183	36	219
Clinical Negative	24	411	435
Column Total	207	447	654 (N)

Table 3: Accuracy Metrics (*P*=8.57e-142<0.05 Significance)

Accuracy Metrics (%)		95% CI
F-Score*	85.9	[81.54, 90.29]
Accuracy	90.8	[88.61, 93.04]
Sensitivity	83.6	[78.65, 88.47]
Specificity	94.5	[92.34, 96.63]
PPV	88.4	[84.04, 92.77]
NPV	91.9	[89.42, 94.47]

*F-score is the harmonic mean of positive predictive value (PPV) and sensitivity. It is a measure of accuracy that takes into account both how sensitive and specific the model is to amblyopia

Table 4: Comparison with other ARF vision screening solutions

	Plusoptix	SPOT	iScreen	GoCheckKids	Kanna
Sensitivity (%)	83	80	75(A), 92(M)*	76(A), 81(M)	83
Specificity (%)	86	85	88(A), 88(M)	67(A), 91(M)	91
Inconclusive (%)	23	4	13(A), 0(M)	6(A), 3(M)	0
ARF Tested	Refractive errors, strabismus	Refractive errors, strabismus	Refractive errors, opacities, ptosis	Refractive errors, gaze error	All ARF
5-Year Cost	$16,250	$25,890	$31,500	$10,140	$349**
Contact	No	No	Yes	No	No
Training	Yes	Yes	Yes	Yes	No

*A:- Automated grading, M: Remote review with Delta Center Crescent, ARF: Amblyopia risk factors. ** Cost of OnePlus 6T as of 03 July 2020 (₹ 26,048.31)

We compared the results of our deep learning based algorithm with other ARF vision screeners such as Plusoptix, SPOT, iScreen, and GoCheckKids, and the results are given in Table 4. Both iScreen and GoCheckKids have automated grading and manual grading using Delta Center Crescent.[17] We find that our automated algorithm is superior to other automated solutions and is comparable with manual grading in GoCheckKids. Overall, our system is capable of detecting more risk factors and is more sensitive and specific than other ARF vision screening systems.[17]

This study addresses the limitations of our previous analysis of the Kanna photoscreener.[19,26] The sample size used for this study is much larger. Furthermore, an age-appropriate population of children with an average age of 7.87 years was recruited in this study increasing the validity of our findings when compared with our initial pilot study.[26] While photoscreening is crucial only up to the age of 5, the age range from 11 months to 18 years was chosen to validate the potential of Kanna in school screening programs and observing that many other studies have used the same age limits for comparing photoscreeners. We also expect that the use of an android-based tool could help complement the current human resources and methods available to screen children for amblyopia even in older children.

However, the sample pool is small for us to make a statistically relevant age-wise analysis. This can be addressed in the subsequent studies.

We also saw further advantages than had been described earlier; particularly we find that our algorithm is highly accurate, sensitive, and specific and can detect individual risk factors with moderate to high correlation. Furthermore, it is of comparable or higher sensitivity and specificity when compared with other ARF vision screening solutions that use specialized hardware or smartphone attachments.

Isometropia ARF prediction showed low accuracy [Table 1]. However, this fallacy is due to the order in which the risk factors are detected in Kanna: ptosis, strabismus, media opacities, anisometropia, and isometropia. In cases of high severity of risk factors (e.g., complete occlusion due to ptosis), it is impossible to predict the remaining risk factors as the iris and red reflex would not be visible. In such situations, the processing is interrupted, the risk factors calculated up till then are shown, and the participant is predicted as positive by Kanna.

Many isometropia cases having high severity of other risk factors (especially anisometropia and strabismus) are predicted as positive for amblyopia without calculating the isometropia risk. Comparing amblyopia prediction in general with the isometropia risk factor gives a more comprehensive view. A total of 107 (82.9%) of the 129 participants with isometropia ARF were predicted as amblyogenic by Kanna.

The vision screener metrics calculated in Table 4 are based on the 2013 AAPOS guidelines, which show an improvement in accuracy to the 2003 AAPOS guidelines for GoCheckKids.[22,27] We expect that using the 2013 AAPOS guidelines would similarly increase our system's predictive accuracy. However, the current model of Kanna as described in Murali *et al.* is based on the 2003 AAPOS criteria that are age agnostic.[19] This was done to simplify risk identification when used in a field survey where knowledge of the date of birth among the target population may not be accurate and may skew the results.

Due to the COVID-19 (coronavirus disease 2019) pandemic, we anticipate that a no-contact tool would be extremely useful for vision screening. As Kanna does not require any hardware add-ons to the smartphone, it is more scalable compared to the available vision screeners that depend on customized hardware. This allows it to be useful also as an at-home amblyopia screening tool encouraging caregivers to seek timely eye care.

Conclusion

We have demonstrated that the "Kanna" photoscreener that uses deep learning and computer vision on photographs of the face is able to reliably detect amblyopia risk factors and could be a scalable and effective option for screening children for amblyopia with minimal resources.

Conflicts of interest

There are no conflicts of interest.

References

1. Holmes JM, Clarke MP. Amblyopia. Lancet 2006;367:1343-51.
2. Fu Z, Hong H, Su Z, Lou B, Pan C-W, Liu H. Global prevalence of amblyopia and disease burden projections through 2040: A systematic review and meta-analysis. Br J Ophthalmol 2020;104:1164-70.

3. Webber AL, Wood J. Amblyopia: Prevalence, natural history, functional effects and treatment. Clin Exp Optom 2005;88:365-75.

4. Daw NW. Critical periods and amblyopia. Arch Ophthalmol 1998;116:502.

5. Donahue SP, Ruben JB. US preventive services task force vision screening recommendations. Pediatrics 2011;127:569-70.

6. Saxena R, Sharma P, Pediatric Ophthalmology Expert Group. National consensus statement regarding pediatric eye examination, refraction, and amblyopia management. Indian J Ophthalmol 2020;68:325-32.

7. Salcido AA, Bradley J, Donahue SP. Predictive value of photoscreening and traditional screening of preschool children. J AAPOS 2005;9:114-20.

8. Wall TC, Marsh-Tootle W, Evans HH, Fargason CA, Ashworth CS, Hardin JM. Compliance with vision-screening guidelines among a national sample of pediatricians. Ambul Pediatr 2002;2:449-55.

9. Mu Y, Bi H, Ekure E, Ding G, Wei N, Hua N, *et al*. Performance of spot photoscreener in detecting amblyopia risk factors in chinese pre-school and school age children attending an eye clinic. Plos One 2016;11:e0149561.

10. Zhang X, Wang J, Li Y, Jiang B. Diagnostic test accuracy of Spot and Plusoptix photoscreeners in detecting amblyogenic risk factors in children: A systemic review and meta-analysis. Ophthalmic Physiol Opt 2019;39:260-71.

11. Kara C, Petriçli IS. Comparison of photoscreening and autorefractive screening for the detection of amblyopia risk factors in children under 3 years of age. J Am Assoc Pediatr Ophthalmol Strabismus 2020;24:20.e1-8.

12. Juan VD, Sanchez I, Ortiz-Toquero S, Martin R. Advantages, limitations, and diagnostic accuracy of photoscreeners in early detection of Amblyopia: A review. Clin Ophthalmol 2016;10:1365-73.

13. Munson MC, Plewman DL, Baumer KM, Henning R, Zahler CT, Kietzman AT, *et al*. Autonomous early detection of eye disease in childhood photographs. Sci Adv 2019;5:eaax6363.

14. Arnold RW, O'neil JW, Cooper KL, Silbert DI, Donahue SP. Evaluation of a smartphone photoscreening app to detect refractive amblyopia risk factors in children aged 1–6 years. Clin Ophthalmol 2018;12:1533-7.

15. Martin SJ, Htoo HE, Hser N, Arnold RW. Performance of two photoscreeners enhanced by protective cases. Clin Ophthalmol 2020;14:1427-35.

16. Peterseim MMW, Rhodes RS, Patel RN, Wilson ME, Edmondson LE, Logan SA, *et al*. Effectiveness of the GoCheck kids vision screener in detecting amblyopia risk factors. Am J Ophthalmol 2018;187:87-91.

17. Arnold RW, Armitage MD. Performance of four new photoscreeners on pediatric patients with high risk amblyopia. J Pediatr Ophthalmol Strabismus 2014;51:46-52.

18. Arnold RW, Arnold AW, Hunt-Smith TT, Grendahl RL, Winkle RK. The positive predictive value of smartphone photoscreening in pediatric practices. J Pediatr Ophthalmol Strabismus 2018;55:393-6.

19. Murali K, Krishna V, Krishna V, Kumari B. Application of deep learning and image processing analysis of photographs for amblyopia screening. Indian J Ophthalmol 2020;68:1407-10.

20. Wood E, Baltrušaitis T, Morency LP, Robinson P, Bulling A. Learning an Appearance-based gaze estimator from one million synthesised images. In Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications. p. 131-8.

21. Bobier WR, Braddick OJ, Eccentric photorefraction. Opt Vis Sci 1985;62:614-20.

22. Donahue SP, Arnold RW, Ruben JB. Preschool vision screening: What should we be detecting and how should we report it? Uniform guidelines for reporting results of preschool vision screening studies. J Am Assoc Pediatr Ophthalmol Strabismus 2003;7:314-6.

23. Oliphant TE. A Guide to NumPy. Vol. 1. USA: Trelgol Publishing; 2006.

24. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, *et al*. SciPy 1.0: Fundamental algorithms for scientific computing in python. Nat Methods 2020;**17:**261-72.

25. McKinney W. Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference. 2010. Vol. 445. p. 51-6.

26. Amitava AK. Commentary: How useful is a deep learning smartphone application for screening for amblyogenic risk factors?. Indian J Ophthalmol 2020;68:1411.

27. Donahue SP, Arthur B, Neely DE, Arnold RW, Silbert D, Ruben JB. Guidelines for automated preschool vision screening: A 10-year, evidence-based update. J Am Assoc Pediatr Ophthalmol Strabismus 2013;17:4-8.