Capstone for Impact Submission | GY2021

Project Title: EyeScreen: Partnering to Develop a Screening Tool for Leukocoria in Ethiopia

Student Name(s): Tochukwu Ndukwe & Sahal Saleh

Advisor Names(s): Dr. Christine Nelson

If this project can be continued by another UMMS student, please include contact information and any other details you would like to share:

Alec Bernard (GY 2022) will continue this project.

Summary (~250-500 words):

Purpose:

Early diagnosis and treatment of retinoblastoma are of paramount importance for a positive clinical outcome. The most common sign of retinoblastoma is leukocoria, or white pupil. Incidence and mortality rates of retinoblastoma is higher in lower-income regions such as Asia and Africa. Effective, easy-to-perform, community-based screening is needed to improve outcomes in these countries. The EyeScreen Android smartphone application works to address this need. The purpose of this study is to examine the potential of the novel use of low-cost technologies - a cell phone application and machine learning - to identify leukocoria.

Design:

Prospective evaluation of using novel technologies - a cell phone application and machine learning - to identify leukocoria.

Participants:

1200 subjects recruited from busy ophthalmology and pediatric clinics in Addis Ababa, Ethiopia.

Methods:

A cell phone application was developed and refined with the feedback from on-site use in Ethiopia. Photos of participant eyes taken at an eye clinic using inexpensive Android smartphones. Photos were reviewed by an ocular oncologist for the pupil color abnormality and were then used to train a ImageNet (Resnet) machine learning model.

Main Outcome Measures:

The performance of the model was measured in terms of sensitivity, specificity, positive predictive value, and negative predictive value.

Results:

Over 4000 images (multiple gaze directions per participant) were collected from Addis Ababa, Ethiopia. 80% of the images were used in training the model, and 20% were reserved for testing. Analyses of the images resulted in the following: sensitivity 87.14%, specificity 73.35%, positive predictive value 0.96, and negative predictive value 0.73.

Conclusions:

EyeScreen is an effective, end-to-end system that uses machine learning algorithms to provide accurate, timely diagnosis of pupil color changes. The cell phone app has shown workability in the target population of Ethiopia and has the potential to serve as an effective screening tool in the areas of the world most affected by delayed retinoblastoma diagnosis. The high performance of the machine learning model with small training datasets can serve as a proof-of-concept for future use of machine learning/AI in ophthalmologic applications.

Methodology:

This study was approved by the University of Michigan Institutional Review Board and by the Institutional Review Board of the St Paul's Hospital Millennium Medical College in Addis Ababa, Ethiopia (HUM00090656). Informed consent was obtained from all participants and all research adhered to the tenets of the Declaration of Helsinki.

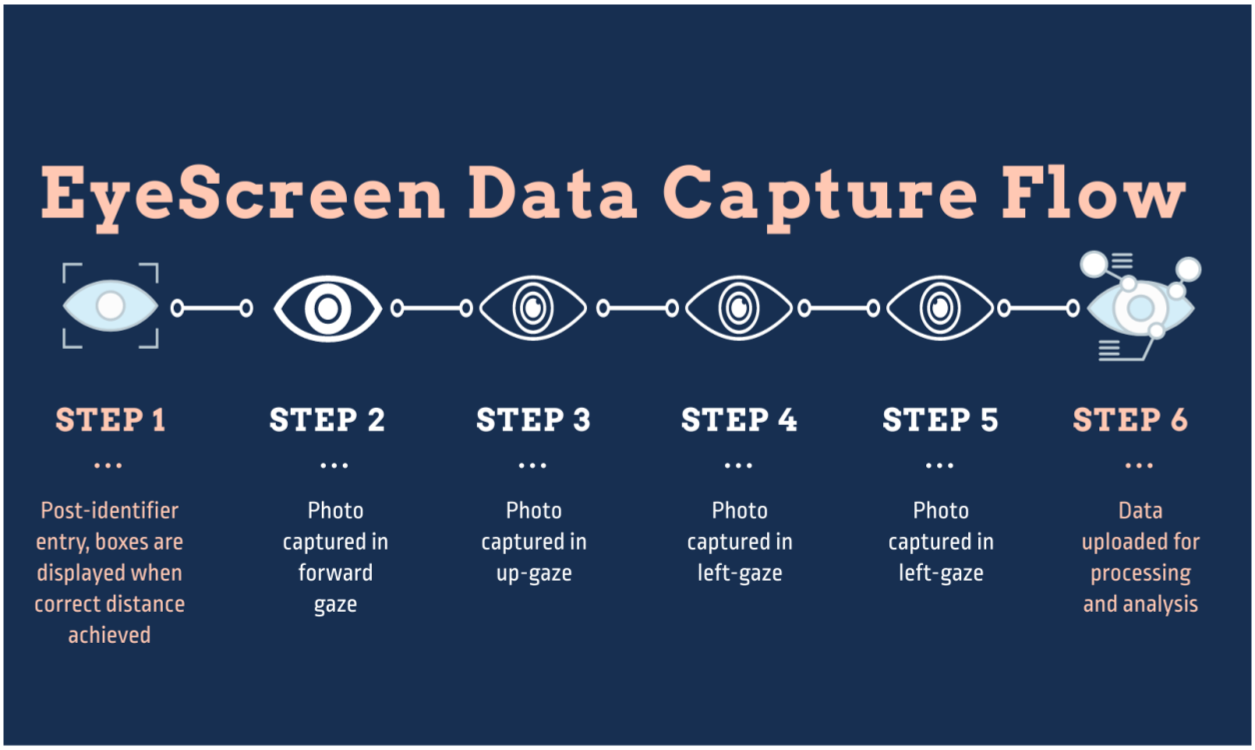
The EyeScreen software is a smartphone application designed for use with Android devices (Google LLC, Mountain View, CA). The EyeScreen application was installed in Google Pixel 3a phones provided to the researchers for the duration of this study. Participants were recruited from St Paul's Hospital Millennium Medical College, a crowded and busy hospital in Addis Ababa, Ethiopia. Researchers were located in the general ophthalmology clinic, pediatric ophthalmology clinic, neonatal ICU, emergency department, and general pediatric clinic.

Consent was obtained from English speaking participants and their parents for underage participants. A local translator was used for patients who spoke Amharic. Patients were seated and lights were dimmed as much as possible (sometimes requiring sheets over the windows). Four different directions of gaze were captured for each participant: up-gaze, left-gaze, right-gaze and center-gaze. Small stuffed animals (beanie babies) were used to assist in directing gaze for younger participants.

The development of EyeScreen was highly iterative and technologically challenging, responding to issues arising in the clinical setting. For example, current smartphone cameras use a preemptive flash technique where bursts of light are sent out before the photo is taken. This action constricts pupils to remove the common "red eye" effect. However, in our situation, the preemptive flash technique is counterproductive as a solid red reflex has a definite impact on subsequent analysis. In order to capture the eye in a maximum naturally dilated (without chemical dilation) state with the red reflex present, we needed to first find the pre-flash setting that was buried deep in the system software. We did find a way to disable the pre-flash light, thereby avoiding the issue of pupillary constriction. Additionally, having the patient in a darker environment assisted in obtaining a clearer view of the red reflex. But, limitations in on-site lighting conditions required still further adjustments to the light-processing balance in the application software. Still further, a lack of reliable wireless or cellular connection required the application to be updated to temporarily store photos until an upload opportunity became available.

To aid the picture taker - and standardize the distance the cell phone was from a participant - EyeScreen displayed boxes around the participant's eyes when they were detected to be the proper distance. The boxes were a sign to the picture taker to then take the picture. The image of the participant's eyes were then displayed on the screen. EyeScreen provided the picture taker with the option to accept the pictures and move to the next direction of gaze, or to retake the image (Figure 1). This process allows for multiple attempts at each gaze if the participant was moving or the image was blurry. The user interface for the Android application displayed helpful tips for optimal photography and the system required no specialized training (Figure 2).

Figure 1: Basic screening steps in use of EyeScreen application



powered by PIKTOCHART

Participants age, race/ethnicity, sex, and presence of ocular conditions were recorded and attached to their images. In the ophthalmology clinics, these ophthalmologic conditions were taken from the patient's chart. All patient information collected was de-identified.

The images underwent automatic image processing within the application to bound the eyes in boxes to preserve participant privacy. The images were then uploaded to a secure server at the University of Michigan. The images were then reviewed by an Ocular Oncologist (**) and classified into simple categories (Normal vs. Abnormal) (Figure 3). These images were then used to fine-tune a pre-trained machine learning model. Our application uses a ImageNet model, specifically: Resnet, an open-source deep-learning network developed for use in image processing applications.10 The ImageNet deep learning model was trained with 80% of the images and tested on the remaining 20%.

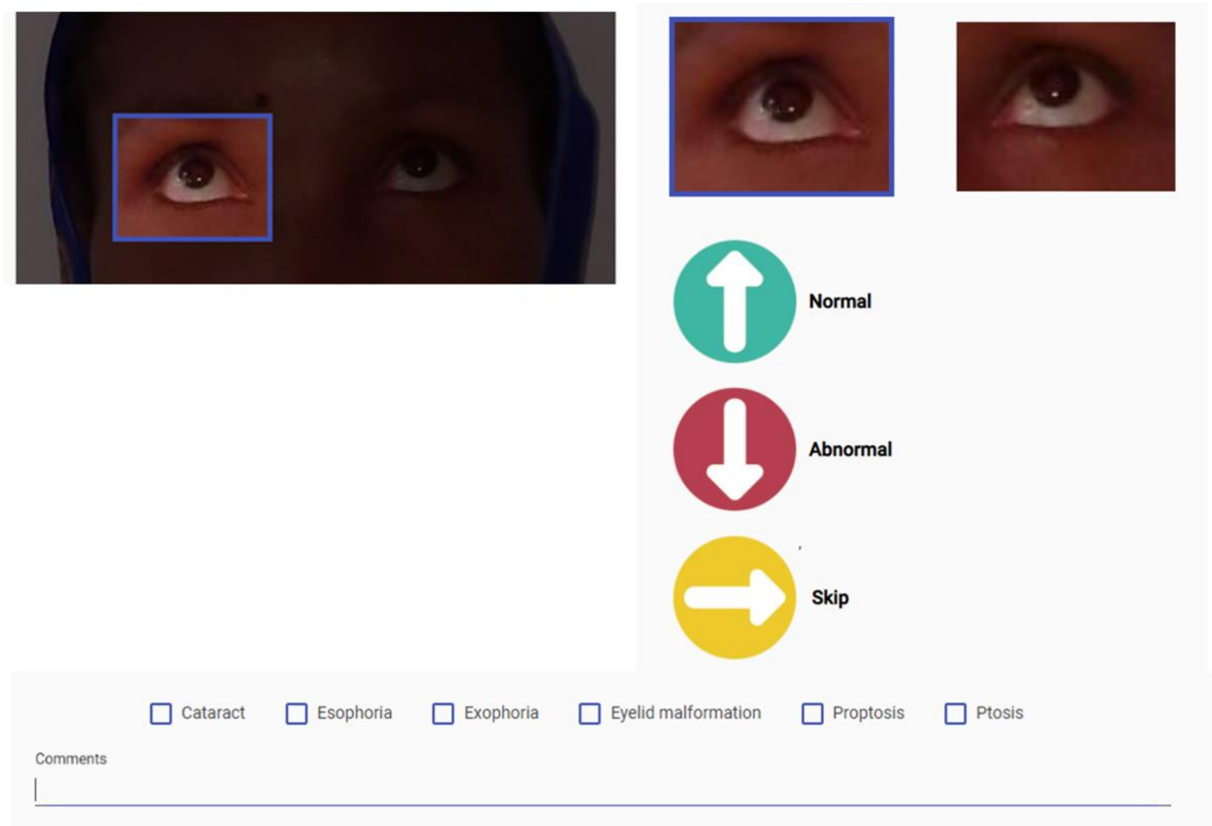


Figure 3: Ophthalmology-facing user interface used to label images for network training.

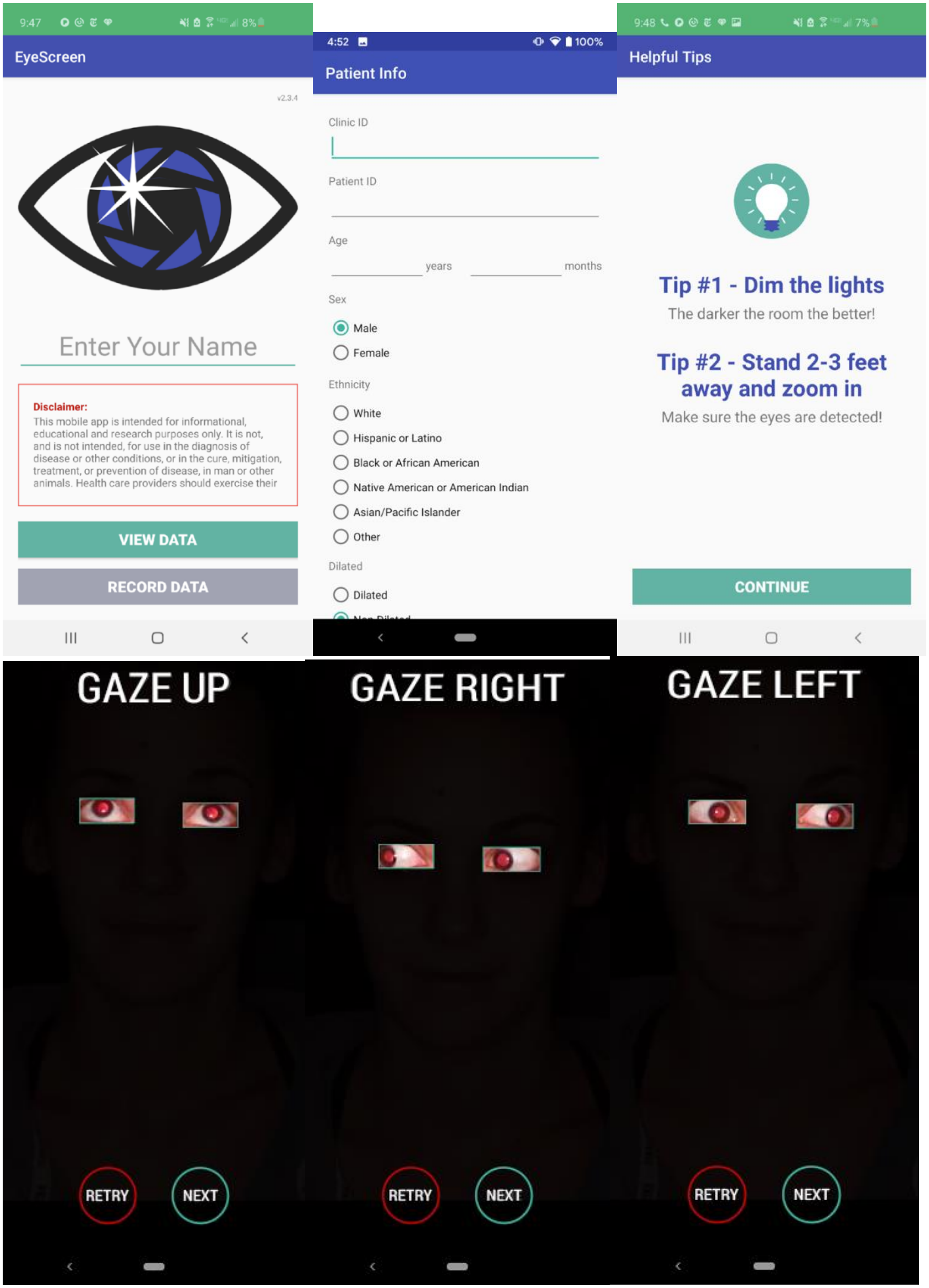
Results:

Over 4000 images of eyes were taken, about four per participant. Table 1 shows demographic information of participants and the distribution of normal red reflex and abnormal red reflex labeled eyes.

	Total Photos Captured (Eye Pairs)	**Eyes Classified as normal red reflex**	**Eyes Classified as abnormal red reflex**
Total Counts	*4356*	*4025*	*331*
Age (mean)	*14.2*	*13.3*	*15.2*
0-2 (count)	*1300*		*116*
3-12 (count)	*2300*		*360*
13-19 (count)	*416*		*44*
20-29 (count)	*144*		*27*
30-89 (count)	*640*		*96*
Sex			
Female	*40%*	*41%*	*33%*
Male	*60%*	*59%*	*67%*

Table 1: Demographics of Participants.

Eighty percent of the images were used in training the model with multiple Resnet training processes completed. The remaining 20% of the images were used to test the accuracy of the model. Sample images used in testing are shown in Figure 4.



EyeScreen
v2.3.4
Enter Your Name
Disclaimer:
This mobile app is intended for informational, educational and research purposes only. It is not, and is not intended, for use in the diagnosis of disease or other conditions, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals. Health care providers should exercise their
VIEW DATA
RECORD DATA
Patient Info
Clinic ID
Patient ID
Age
years
months
Sex
Male
Female
Ethnicity
White
Hispanic or Latino
Black or African American
Native American or American Indian
Asian/Pacific Islander
Other
Dilated
Dilated
Helpful Tips
Tip #1 - Dim the lights
The darker the room the better!
Tip #2 - Stand 2-3 feet away and zoom in
Make sure the eyes are detected!
CONTINUE
GAZE UP
RETRY
NEXT
GAZE RIGHT
RETRY
NEXT
GAZE LEFT
RETRY
NEXT

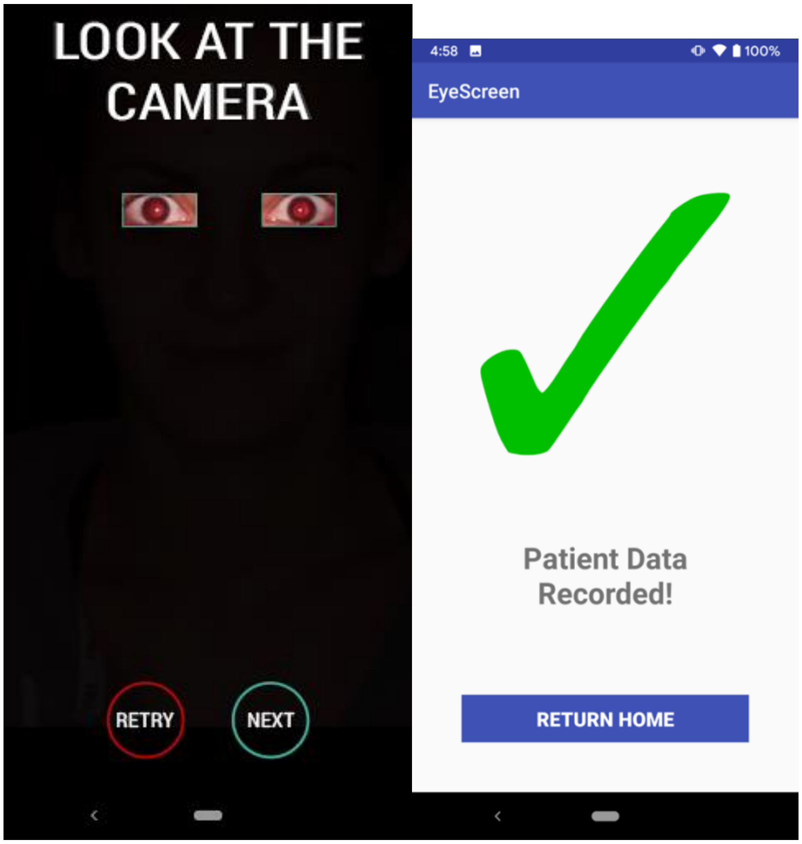
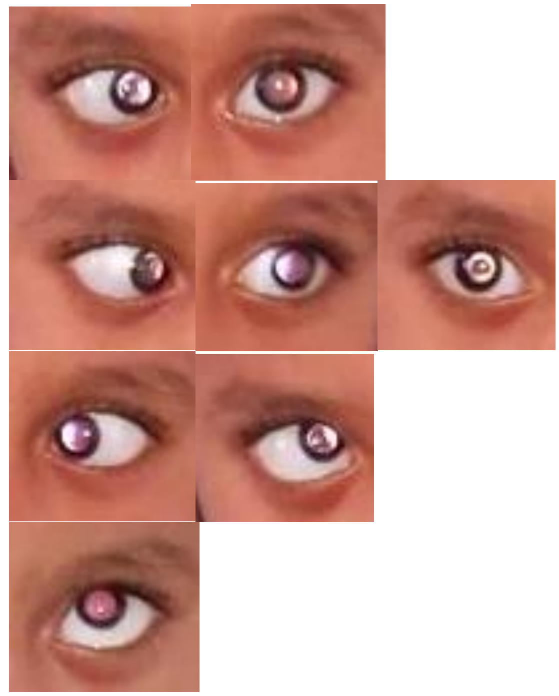
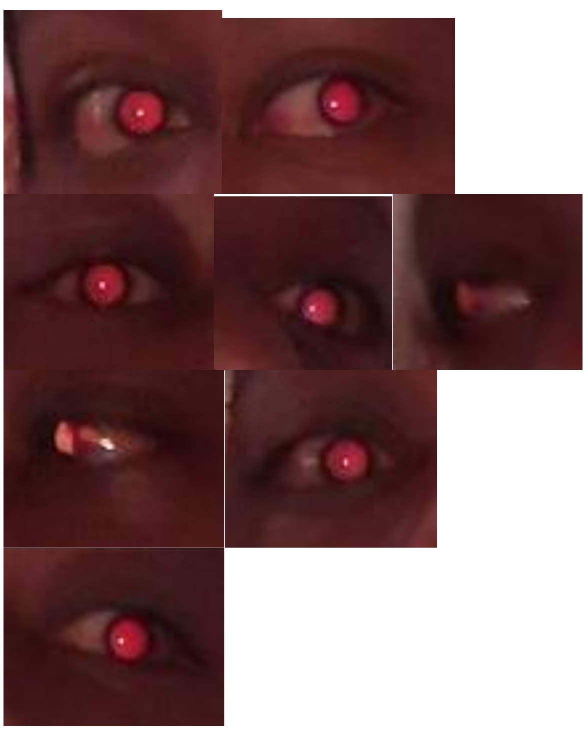
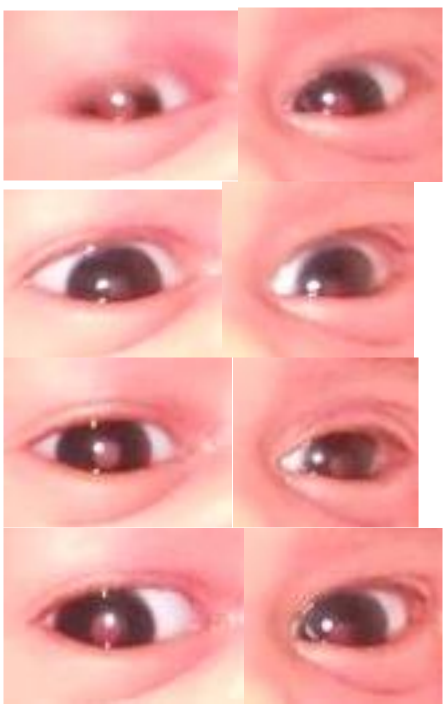
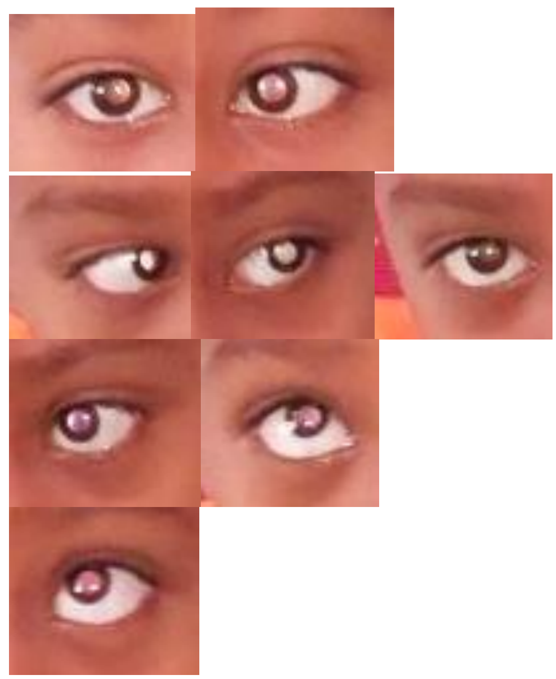
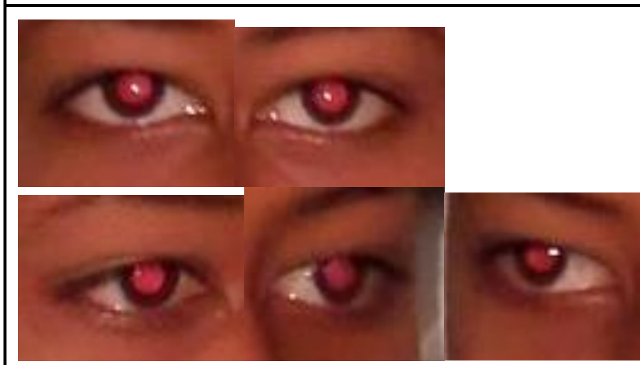
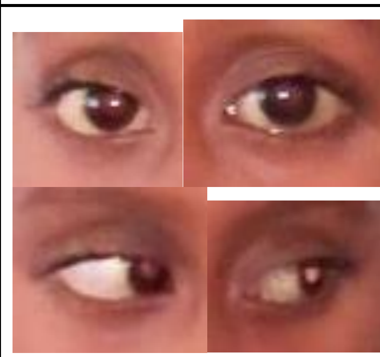


Figure 2. Screenshots of the user interface for EyeScreen before taking photographs.

Normal red reflex	**Abnormal pupil reflex**
 Participant A	 Participant F
 Participant B	 Participant G

 Participant C	 Participant H
 Participant D	 Participant I
	

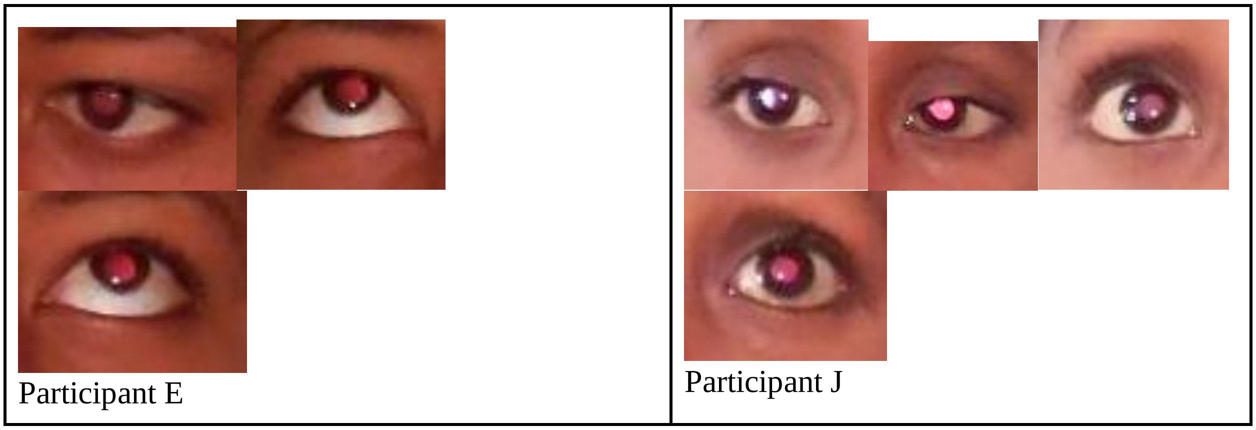


Figure 4: Sample of images labeled normal red reflex and abnormal red reflex in model training and testing.

In the 869 photos of test eyes (210 normal red reflex, 30 abnormal red reflex), EyeScreen was 95% accurate in detection of abnormal vs normal red reflex. The initial sensitivity after training on our dataset was 87.14% with a specificity of 73.35%. The positive predictive value was 0.96 and the negative predictive value was 0.73 in preliminary trials with the test set of images. The Receiver Operating Characteristic (ROC) curve for this dataset is shown in Figure 5.

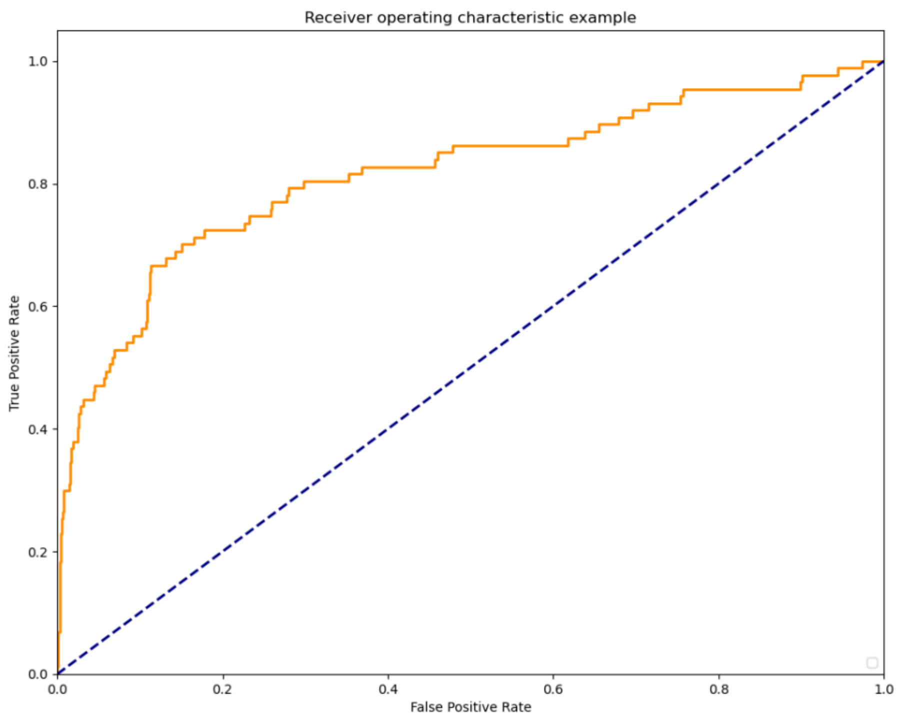


Figure 5: Receiver Operating Characteristic (ROC) curve for EyeScreen testing dataset.

Conclusion (~250-500 words):

Compared to high-income countries, children with retinoblastoma from low-income countries present at an older age with more advanced disease and a higher rate of metastasis.3 Patients from high-income countries were diagnosed at a median age of 14.1 months, with less than 1% of patients having extraocular retinoblastoma and less than 1% having metastasis. In contrast, patients from low-income countries were diagnosed at a median age of 30.5 months, with 49.1% having extraocular retinoblastoma and 18.9% having metastasis. Clearly, there is a significant delay in the recognition of retinoblastoma in low-income countries, when compared with recognition of retinoblastoma in high-income countries. A review of presentations of retinoblastoma in Ethiopia found that the most common presenting sign was proptosis, signifying late presentation and late referral patterns for these patients.11 Considering 85% of global retinoblastoma cases are from the low-income countries, there is an unmet need for effective, easy, and low-cost screening tools for retinoblastoma.

Leukocoria or white pupil is the most common presenting symptom in 63% of cases, globally. In a meta-analysis, Subhi et al. reported that estimated sensitivity of abnormal red reflex testing for detecting ocular pathology was 7.5% and specificity was 97.5%. The positive predictive value was 53% and negative predictive value was 74%.12 Leukocoria or white pupil can be detected via flash photography. There are two smartphone-based screening applications that evaluate the pupil color; both have significant drawbacks, as they are both based on the more costly iOS system. Android phones - typically available at much lower price points - make up 85% of the cellular phone sector while iOS phones make up 12% as of December 2020.13 There is a need, then, for an Android phone-based application such as EyeScreen. Given the widespread use of Android smartphones in Sub-Saharan Africa, and other resource-limited settings, a leukocoria-detecting application would be accessible for users in these areas.

CRADLE, an Android-based application, examines existing photos on the user's device and as such, would not be effective as a community-based screening tool.14 Another Android-based application, MDEyeCare, was assessed in a small study of 28 patients; it requires relatively standard conditions for photos which may be difficult to achieve in a real-world setting. EyeScreen is a significant improvement since it addresses the limitations of the other two Android applications.

In this study, in a resource-limited setting and under real-world conditions, we have demonstrated the clinical utility of the EyeScren application. EyeScreen obtained sensitivity and specificity ratings significantly higher than current red reflex testing. This sensitivity and specificity will also continue to improve as additional photographs and populations are added into the training dataset. In addition to the use of low-cost smartphones, the use of low-resource demand machine learning technology also made EyeScreen an effective tool. Additionally, we have demonstrated utility in a population subject to typical increased difficulty in screening. Children with darker fundus pigmentation may have abnormal reflex testing results because an increased amount of melanin pigment can give a duller red reflex. This study has demonstrated that the EyeScreen application, with its low-cost, efficient technologies, can be an effective tool in a resource-limited setting.

Current community-based screening protocols have poor sensitivity or require training and do not reach many of the most vulnerable populations.15 Development of effective, free, and simple-to-use applications like EyeScreen could allow increased screening of the populations in community-based settings, allowing earlier referral and treatment.

In this study, we successfully developed an effective end-to-end system using transfer learning algorithms to provide an accurate and timely diagnosis of a key presentation of retinoblastoma. Work is ongoing to allow detection of additional ocular pathologies within the application and increase its sensitivity for global use.

Reflection/Impact Statement:

You may use the following questions to guide your reflection:

1. How did the process of conducting this research confront any limitations of your prior thinking?
2. Who could potentially benefit from this CFI project over different timescales and how?
3. What actions will you take afterwards to continue the momentum of this project, and maximize the likelihood of the identified benefits being achieved?
4. What advice would you give to another student completing their CFI?

My roles during this trip included being a learner, collecting quality data, and performing early analysis of those data during the trip. This was my first delve into the world of ophthalmology from a research perspective, and it gave me solid footing going forward as it relates to the execution of research in low-resource settings. This experience has allowed me to compare and contrast the differences in our patient populations, approaches to care, and cultural engagement. Furthermore, the collaboration between the Kellogg Center and the ophthalmologists of St. Paul’s Millennium Medical College is a great example of a sustainable educational endeavor, and I hope that I can learn from this collaborative effort in any future travels.

This project fundamentally aims to improve the vision, lives, and health of those most vulnerable – children in low resource settings. Our phone-application provided evidence that there are affordable, sustainable approach to improving a specific screening protocol for retinoblastoma. Importantly, this project could be expanded such that we could screen for other eye diseases including strabismus or ptosis, or apply it in other fields as well.

We’ve since submitted a transcript of this project and we hope to get published soon. Beyond, we anticipate expanding from screening for red reflex and leukocoria to other eye diseases. We also hope that once the pandemic is under adequate control, we can return to St. Paul’s to implement the screening tool on a wide-scale basis and to truly make an impact on people’s lives, vision, and well-being.

Citations

1. Kivela T. The epidemiological challenge of the most frequent eye cancer: retinoblastoma, an issue of birth and death. *British Journal of Ophthalmology*. 2009;93(9):1129-1131. doi:10.1136/bjo.2008.150292
2. Dimaras H, Kimani K, Dimba EA, et al. Retinoblastoma. *The Lancet*. 2012;379(9824):1436-1446. doi:10.1016/S0140-6736(11)61137-9
3. Fabian ID, Abdallah E, Abdullahi SU, et al. Global Retinoblastoma Presentation and Analysis by National Income Level. *JAMA Oncol*. 2020;6(5):685. doi:10.1001/jamaoncol.2019.6716
4. Leander C, Fu LC, Peña A, et al. Impact of an education program on late diagnosis of retinoblastoma in Honduras. Pediatric Blood & Cancer. 2007;49(6):817-819. doi:https://doi.org/10.1002/pbc.21052
5. Abramson DH, Beaverson K, Sangani P, et al. Screening for Retinoblastoma: Presenting Signs as Prognosticators of Patient and Ocular Survival. *PEDIATRICS*. 2003;112(6):1248-1255. doi:10.1542/peds.112.6.1248
6. Shields CL, Gorry T, Shields JA. Outcome of Eyes With Unilateral Sporadic Retinoblastoma Based on the Initial External Findings by the Family and the Pediatrician. *J Pediatr Ophthamol Strabismus*. 2004;41(3):143-149. doi:10.3928/0191-3913-20040501-06
7. Canturk S, Qaddoumi I, Khetan V, et al. Survival of retinoblastoma in less-developed countries impact of socioeconomic and health-related indicators. *British Journal of Ophthalmology*. 2010;94(11):1432-1436. doi:10.1136/bjo.2009.168062
8. Li J, Coats DK, Fung D, Smith EO, Paysse E. The Detection of Simulated Retinoblastoma by Using Red-Reflex Testing. *PEDIATRICS*. 2010;126(1):e202-e207. doi:10.1542/peds.2009-0882
9. Nyamori JM, Kimani K, Njuguna MW, Dimaras H. The incidence and distribution of retinoblastoma in Kenya. *Br J Ophthalmol*. 2012;96(1):141-143. doi:10.1136/bjophthalmol-2011-300739
10. Raghu M, Zhang C, Kleinberg J, Bengio S. Transfusion: Understanding Transfer Learning for Medical Imaging. *Advances in Neural Information Processing Systems*. 2019;32:3347-3357.
11. Shifa JZ, Gezmu AM. Presenting signs of retinoblastoma at a tertiary level teaching hospital in Ethiopia. *Pan Afr Med J*. 2017;28. doi:10.11604/pamj.2017.28.66.11199
12. Subhi Y, Schmidt DC, Al-Bakri M, Bach-Holm D, Kessel L. Diagnostic Test Accuracy of the Red Reflex Test for Ocular Pathology in Infants: A Meta-analysis. *JAMA Ophthalmol*. Published online November 12, 2020. doi:10.1001/jamaophthalmol.2020.4854
13. Mobile Operating System Market Share in Africa - December 2020. Published online January 10, 2021. Accessed January 10, 2021. https://gs.statcounter.com/os-market-share/mobile/africa
14. Khedekar A, Devarajan B, Ramasamy K, Muthukkaruppan V, Kim U. Smartphone-based application improves the detection of retinoblastoma. *Eye*. 2019;33(6):896-901. doi:10.1038/s41433-018-0333-7
15. Mndeme FG, Mmbaga BT, Kim MJ, et al. Red reflex examination in reproductive and child health clinics for early detection of paediatric cataract and ocular media disorders: cross-sectional diagnostic accuracy and feasibility studies from Kilimanjaro, Tanzania. *Eye*. Published online June 16, 2020. doi:10.1038/s41433-020-1019-5