HIGH SCHOOL EDITION
Journal of Student Research

Volume 10 Issue 3 (2021)

AI Democratization in Optometry: Developing a Prototype with Azure Cognitive Services Platform

Shaun Baek[1], Ryan Johnson[1#], Claire Saunders[2#], Debora Lee Chen[3#], Katherine Lai[3#]

[1]Portola High School, Irvine, CA, USA
[2]California Institute of Technology, Pasadena, CA, USA
[3]University of California, Berkeley School of Optometry, Berkeley, CA, USA
[#]Advisor

ABSTRACT

As Artificial Intelligence (AI) technology advances, it is used in almost every aspect of our lives. However, AI is still complicated to implement without help from computer engineers. In the healthcare field, knowledge of medical and computer knowledge is necessary to create AI-based medical systems. Close cooperation between medical experts and computer experts is essential. For this reason, even if there has been a continuous effort to apply AI into the medical field, it has yet to be universalized. In particular, in the field of optometry and ophthalmology, more complex technology is required than in other medical fields because it is necessary to analyze an eye image to diagnose a disease. Therefore, this study explores the possibility for medical professionals with little computer knowledge in the field of ophthalmology to develop an AI-based diagnostic system without the help of computer engineers. In addition, it explores not only the possibilities but also the diagnostic accuracy of the developed system. Our results show that the diagnostic system discriminates against five common eye diseases to some extent. This study explores whether AI democratization is possible even in the field of ophthalmology that requires advanced skills and knowledge.

Introduction

Medical services are essential parts of wellness and well-being. However, many people do not receive necessary health care due to the lack of accessibility and affordability. This lack of preventative care results in a more significant number of emergency visits. Experts in the medical and technology fields are acutely aware of this issue and are developing innovative ways to integrate information technology with healthcare. Through this, they hope to create a more affordable and accessible experience for everyone. One of the ways they are doing this is through telemedicine.

Telemedicine includes but is not limited to virtual consultations with medical professionals. More advanced systems using AI remotely diagnose conditions using artificially trained programs. These new technologies improve the accessibility issue; however, telemedicine continues to be expensive to use. One of the fields that have been proven to be the least developed is the optometry field. Tele-optometric devices are not readily available to the public, especially in remote or underdeveloped areas. This demand for tele-optometric devices results in an increased cost in preventative care, which is often as expensive as in-person care (Hulsey, 2020). Preventative measures are essential to ensure that everyone can take appropriate measures to maintain their eye health. In order to develop intelligent tele-optometric systems, a large amount of training data is required. A lack of available datasets [and technically Health Insurance Portability and Accountability Act (HIPAA) regulations] makes it difficult for medical practitioners to develop these systems. In this paper, we develop an intelligent system for diagnosing eye diseases using Microsoft Azure's Cognitive Service platform and open-access datasets. We explore how useful this system can be developed for diagnosing eye diseases. For AI to be widely used, more people should be able to use and benefit from AI technologies without any specialized knowledge. Recently, this has been known as AI Democratization (Garvey, 2018). In this study, we intend to verify this possibility in the field of optometry.

AI in Optometry

Optometrists recommend individuals between the ages of 20 and 50 receive eye check-ups every two to five years. As people age or show deteriorating eye health, the number of recommended eye examinations increases. Despite these recommendations, a large proportion of the population lacks proper optometry care. Reasons for this include economics (39.8%), refusal for care (35%), and scheduling issues (4.5%) (Hendrick, 2011). The out-of-pocket cost per optometry visit is over 200 US Dollars. This cost results in individuals only seeking care for urgent matters (Taylor et al., 2004). Over 26.1 million people suffer from a lack of health coverage, excluding those that exclusively own accidental insurance (Bureau, 2020). The novel Coronavirus (COVID-19) pandemic significantly increased the number of uninsured individuals with estimates as high as 6.2 million (Bivens & Zipperer, 2020). The American Optometric Association (AOA) recommends regular eye examinations—including presenting and best-corrected visual acuity, visual field testing, and comprehensive eye examination with dilation—to detect early signs of conditions or diseases that are not even related to the eyes (American Optometric Association). Factors including excessive screen time, poor diet, and lack of sleep aggregate underlying eye conditions (Primera Eye Care, 2020).

Telemedicine was initially met with skepticism. Over the past decade, the significant increase in its research improved trust in telemedicine and its subdisciplines, including tele-optometry. Tele-optometry is a distanced eye health service that utilizes medical technology to make eye care more accessible (Optix Family Eyecare, 2020). This technology enables eye doctors to screen remotely for certain eye diseases through video conferencing technology. However, the progression in optometry is slower than in other medical fields because of the expensive equipment for eye examinations (Sharma et al., 2020).

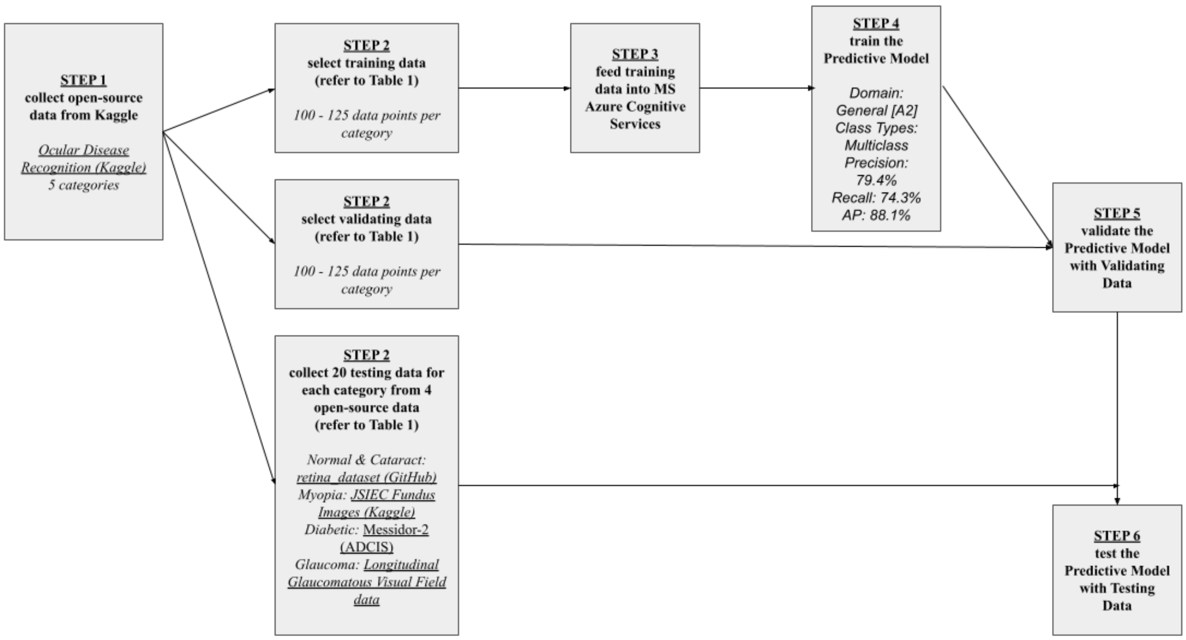
The first medical system that applied artificial intelligence was MYCIN, developed by Stanford University in 1972 (Copeland, 2020). This system treated blood infections and diagnosed patients through self-reported symptoms and medical test results (Shortliffe, 1977). It provided diagnostic results and recommendations with detailed explanations for patients to understand. It has been proven to operate as well or better than medical specialists in detecting blood infections. However, it was not widely used due to the hefty cost and lack of confidence in these systems from doctors and the general public.

As AI continues to advance and solve increasingly complex problems, there has been an increased application in optometry, where visual information is critical for diagnostic purposes. Despite the continued lack of accessibility and coverage, these AI-based optometric devices have been very effective in notifying patients before experiencing severe vision loss.

Currently, AI applications are being developed for specific eye diseases—specifically diabetic retinopathy (DR), glaucoma, and retinopathy of prematurity (ROP)—which are the leading causes of vision loss (Noronha & Nayak, 2012). Many believe that AI would be useful in quickly identifying patients with preventable vision loss. This is especially crucial in developing countries where access to medical doctors is difficult due to the lack of medically trained individuals. People in these regions can use alternatives such as smartphones and mobile applications, which are capable of diagnosing diseases without large, expensive machines (Mrutyunjaya & Raga, 2020). However, it has been challenging for medical professionals and engineers to expand their patients due to the lack of trust in AI systems (Guo et al., 2019).

Developing & Testing A Prototype

This research aims to determine the accuracy and usability of an eye diagnostic tool using open-source development tools and datasets. In this study, the Azure Cognitive Services Platform, which can be easily used without prior knowledge in AI algorithms or programming, is used as the system development tool. In addition, the data for training the system is also collected from several public data repositories. [Figure 1] summarizes the detailed research process.



*Ocular Disease Recognition (Kaggle) (Shanggong Medical Technology Co., Ltd., 2020)
*retina_dataset (GitHub) (Chen)
*JSIEC Fundus Images (Kaggle) (Joint Shantou International Eye Centre, 2019)
*Messidor-2 (ADCIS) (Patry, 2021)
*Longitudinal Glaucomatous Visual Field data (Rotterdam Ophthalmic Data Repository)
*General [A2] (Microsoft Azure, 2020)

Figure 1. Research Process

Step 1: Collect Open-Source Data from Kaggle

Open-source data of fundus images of eyes is needed to train and validate a predictive model. The Ocular Disease Recognition dataset from Kaggle is used to train and validate the diagnostic tool (Shanggong Medical Technology Co., Ltd., 2020). This real-life dataset of patient information is collected by Shanggong Medical Technology Co. Ltd. The data comes from different hospitals and medical centers in China and has five categories: Normal, Diabetic, Glaucoma, Cataract, and Myopia. Diabetic, or Diabetic Retinopathy, is a diabetic complication that damages the blood vessels of the back of an eye (Boyd, 2021a). Glaucoma is caused by high pressure in front of an eye, leading to optic nerve damage (Boyd, 2021b). Cataract is when the lens of the eye becomes cloudy or blurry due to a lens' damage (Boyd, 2021c). Myopia, or nearsightedness, is a refractive error that can be inherited by family members (Turbert, 2021).

Step 2: Select Training/Validating/Testing Data

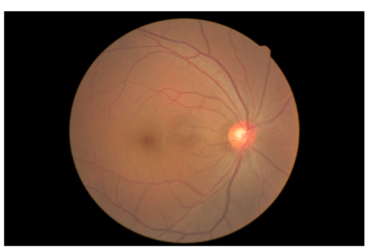
The collected dataset is cleaned, filtered, and sorted using Google Sheets and Python. The data points with indications of multiple diseases will be removed when AI is trained. The data points are randomly filtered to minimize bias and organized into five categories. Each category has 200 to 250 data points for training (from the Ocular Disease Recognition dataset), randomly selected 20 data points for validation (from the Ocular Disease Recognition dataset) that

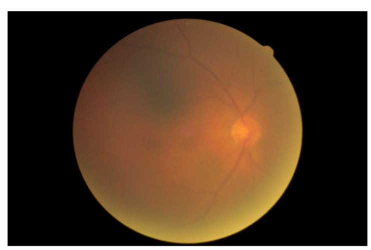
HIGH SCHOOL EDITION
Journal of Student Research
Volume 10 Issue 3 (2021)

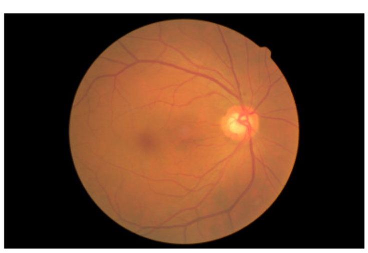
have not been selected for training, and 10 data points as evaluation points (from four datasets used to test the model). In order to increase accuracy in our results for the predictive system, the evaluation data used for testing was collected from four completely different open-source data sites from the data used for training and validation. While this can be done with only Google Sheets, Python helps organize the fundus images of eyes into their corresponding categories. [Table 1] summarizes information about data sources for training, validating, and testing. In addition, [Figure 2] shows examples of eye fundus images.

Table 1. Outline of Training, Validating, and Testing Data

	Training Data	Validating Data	Testing Data
N (Normal)	242 images (https://www.kaggle.com/linchundan/fundusimage1000)	20 images (https://www.kaggle.com/linchundan/fundusimage1000)	10 images (https://github.com/yiweichen04/retina_dataset)
C (Cataract)	136 images (https://www.kaggle.com/linchundan/fundusimage1000)	20 images (https://www.kaggle.com/linchundan/fundusimage1000)	10 images (https://github.com/yiweichen04/retina_dataset)
D (Diabetic)	225 images (https://www.kaggle.com/linchundan/fundusimage1000)	20 images (https://www.kaggle.com/linchundan/fundusimage1000)	10 images (http://www.rodrep.com/data-sets.html)
G (Glaucoma)	134 images (https://www.kaggle.com/linchundan/fundusimage1000)	20 images (https://www.kaggle.com/linchundan/fundusimage1000)	10 images (http://www.rodrep.com/data-sets.html)
M (Myopia)	130 images (https://www.kaggle.com/linchundan/fundusimage1000)	20 images (https://www.kaggle.com/linchundan/fundusimage1000)	10 images (https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k)


Normal
(Joint Shantou International Eye Centre, 2019)


Cataract
(Joint Shantou International Eye Centre, 2019)


Glaucoma
(Joint Shantou International Eye Centre, 2019)

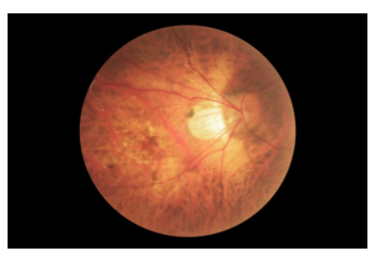

Myopia
(Joint Shantou International Eye Centre, 2019)

Figure 2. Examples of Trained, Validated, and Tested Fundus Images

Step 3: Feed Training Data into the Microsoft Azure Cognitive Services

The organized data points are inputted into the Azure Cognitive Services platform and tagged according to their categories (refer to Figure 3). The Azure Cognitive Service is an AI service platform in which anyone can easily apply AI technology to their work if adequately refined training data is provided.

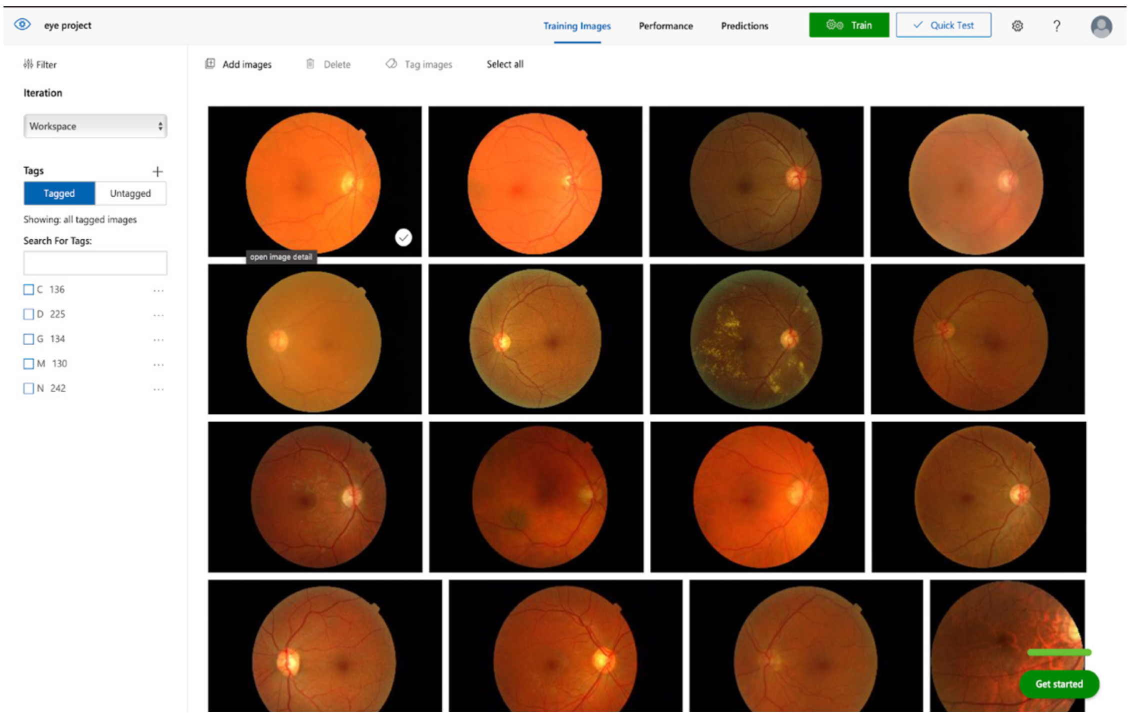


Figure 3. Interface of Azure Cognitive Services Custom Vision

Step 4: Train the Predictive Model

After inputting all data points and their categories, we create and train a predictive model under the General [A2] domain and Multiclass classification type. The General [A2] domain is a lightweight model optimized for better accuracy with fast inference times (Microsoft Azure, 2020). The Multiclass classification is a type that has a single tag per input (Microsoft Azure, 2020).

[Figure 4] shows the training results. The results come out as 79.4% Precision, 74.3% Recall, and 88.1% AP ratings. The Probability Threshold is the minimum probability score for a prediction to be valid when calculating precision and recall (Microsoft Azure). The Precision indicates the chances to output the correct disease when asked to predict. The Recall indicates the chances of finding the right disease type. The AP (also known as F1-score) summarizes the model's performance, considering the Precision and Recall at different Probability Thresholds.

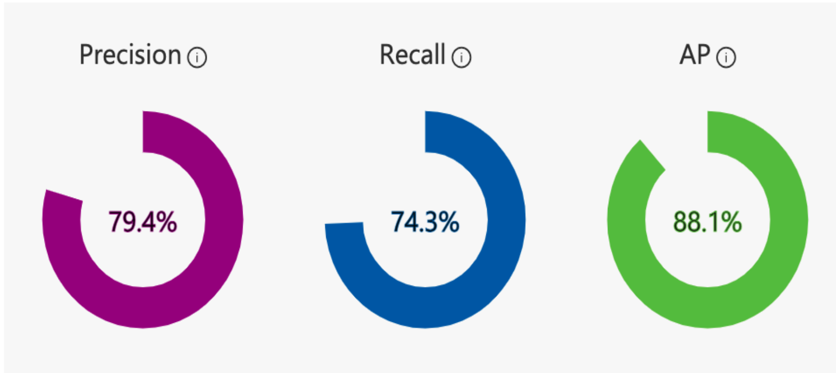


Figure 4. Training Performance from Azure Cognitive Services

Step 5: Validate the Predictive Model with Validating Data

In order to verify the system, the predictability of the system is evaluated with data that has not been used as the training data (refer to Table 1). [Table 2] is the result of validating the trained model. The predictive model uses 20 validation data points from each category to ensure it gets similar results to their corresponding categories. Through these tests, a confusion matrix is created to visualize the predictive model's performance (Mohajon, 2021). From [Table 2], Myopia (M) has the best performance because of the greatest F1-score of 95.24%, while Glaucoma (G) has a worse performance of 70.27%. Despite these differences in performance, the overall average F1-score of 77.91% results in a similar performance outlined by the predictive model in [Figure 2] of 88.1%.

Table *2*. Predictive Model Performance with Validation Data Points

Confusion Matrix		Actual Results					
		N	C	D	G	M	
	N	18	2	5	5		
	C	1	14		1		
	D	1		13	1		
	G		2	2	13		
Predicted Results	M		2			20	Average
	Precision	60.00%	87.50%	86.67%	76.47%	90.91%	80.31%
	Recall	90.00%	70.00%	65.00%	65.00%	100.00%	78.00%
	F1-score	72.00%	77.78%	74.29%	70.27%	95.24%	77.91%

Step 6: Test the Predictive Model

After validating our predictive model, we need to test the model as a next step. The data used for the testing is selected from completely different sources than the data used for training and validation. We used four different sources: Normal and Cataract use a retina dataset collected by Chen on GitHub, Myopia uses a dataset of 1000 Fundus images collected by the Joint Shantou International Eye Centre (2019). Diabetic uses the Messidor-2 dataset collected by A Team of Imaging Experts (ADCIS) (Patry, 2021). Glaucoma uses the Longitudinal Glaucomatous Visual Field Data collected by the Rotterdam Ophthalmic Data Repository.

 Volume 10 Issue 3 (2021)

The 10 evaluation data points for each category are tested through the predictive model (refer to Figure 5). The results for each category are shown in Table 3. The categories are chosen by the top three highest average percentage rating out of the five possible diseases. The percentage indicates the chances of each tag being similar to the classified data points: the higher the percentage, the higher the likelihood the predictive model's indicated disease corresponds to the image input. The model accurately identifies eye images associated with four other eye diseases except for Glaucoma. Glaucoma eye images show a higher probability of being misdiagnosed as a normal eye with a difference of 13.55%.

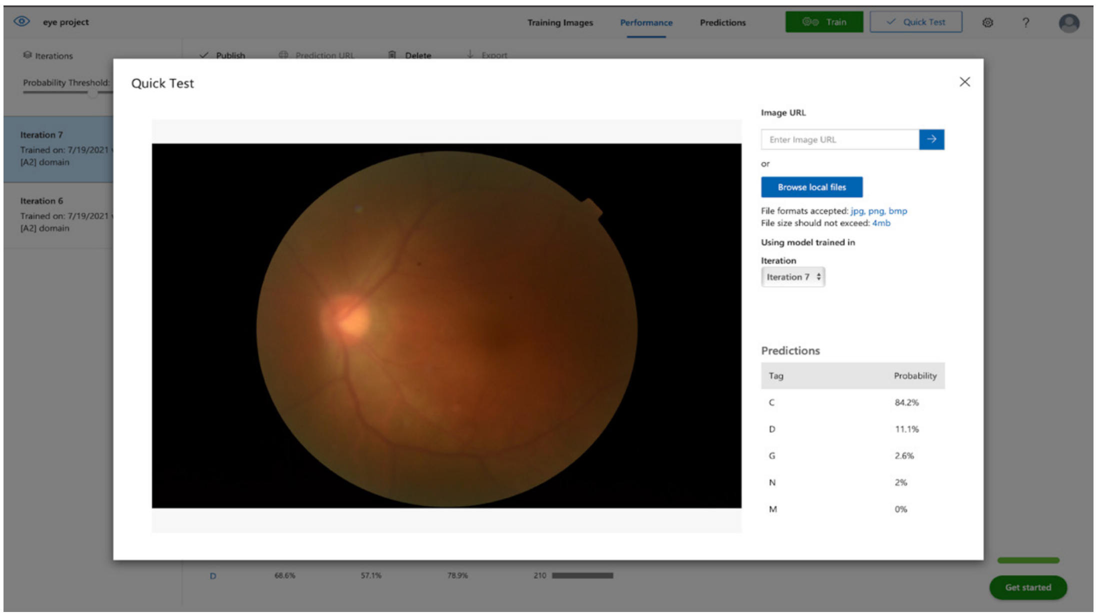


Figure 5. Testing Interface on Azure Cognitive Services

Table *3*. Predictive Model Performance with Evaluation Data Points

Cataract					Normal			
Test Case #	C (%)	G (%)	N (%)		Test Case #	N (%)	G (%)	D (%)
1	87.2	3.9	7.6		1	88.3	3.1	8.4
2	84.6	13.2	1.9		2	88	2.5	9.3
3	76.7	16.6	5		3	87.9	5.2	5.5
4	39.9	23.2	33.5		4	77.1	3.2	19.5
5	38.6	36.4	11		5	75.5	3.9	20.3
6	34.1	9.1	32.7		6	69.1	1.3	29.1
7	30.2	52.4	12.7		7	65.7	1.7	32.4
8	27.5	68.9	1.5		8	64.7	14.4	20.5
9	5.2	8.6	24.5		9	58.8	2.5	38.3
10	3	7	82.2		10	54.7	40.9	4.2
Average:	42.7	23.93	21.26		Average:	72.98	7.87	18.75

Myopia			
Test Case #	M (%)	D (%)	G (%)
1	99.9	0	0
2	99.7	0.2	0
3	99.4	0.1	0.3
4	99.4	0.2	0.2
5	98.7	0.2	0.5
6	95	1.3	1.6
7	94.5	0.7	0.7
8	93.2	4.1	1.4
9	92	6.2	0
10	87.1	2.5	6.5
Average:	95.89	1.55	1.12

Diabetic			
Test Case #	N (%)	D (%)	C (%)
1	0.1	99.7	0
2	17.4	78.7	0
3	24.4	74	0.2
4	31.7	65.7	0
5	31.1	65.6	0
6	15.3	63.8	16.8
7	37.4	59.9	0
8	50.6	48.3	0.5
9	60.9	36.8	1.1
10	41.5	34.6	0.6
Average:	31.04	62.71	1.92

Glaucoma			
Test Case #	G (%)	N (%)	D (%)
1	88.5	11	0.2
2	67.2	28.7	0
3	48	49.3	2.3
4	42.3	56.5	0.7
5	34.4	64.1	1.3
6	33	55.9	9
7	30.9	48.5	19.7
8	28.3	69	2.4
9	21.3	68	7.8
10	9.6	88	2.2
Average:	40.35	53.9	4.56

Legend
N - Normal
C - Cataract
D - Diabetic
G - Glaucoma
M - Myopia

Conclusion

In this study, we examined how AI can transform diagnostics in tele-optometry with open-source tools. Our results show that for Myopia that AI performed well with a 95.89% accuracy, while we saw for Glaucoma a much higher probability of receiving a misdiagnosis with a 40.35% accuracy. The eye disease diagnosis system that we developed using the Azure Cognitive Service and open data sources empirically verified that it can discriminate five eye diseases to some extent, although there is a slight difference in accuracy.

Currently, there are tools that allow the general public to easily develop AI systems, and the data to train the AI system are accessible (but limited) to the general public. There have been attempts to replace human doctors by developing an AI system in the medical field for many years. However, despite the fact that advanced information and communication technology can be used as it is now, it is thought that it has a long way to become universal. Therefore, in this study, we explore how useful it is to predict eye diseases using Azure Cognitive Services and open datasets. By demonstrating that the AI tool developed by a non-expert makes a diagnosis as accurate as a human doctor, this study shows that human doctors can develop and use tools to help diagnose their patients without any help from engineers or having computer knowledge. It is expected that trust in AI technology will increase as medical practitioners, including doctors and nurses, develop and use AI tools by themselves. In the future, one of the first steps is to amass more eye fundus images to build bigger and better training datasets. These datasets will allow AI systems to be faster and more accurate, which may allow for accurate diagnostics with just a mobile application. With this mobile application, we can create a network of eye doctors to facilitate and review each user's diagnosis to retain accuracy and credibility and continuously retrain the system with user inputs.

Limitations

Study limitations include an inability to determine multiple diseases, bias in the training set, and limited datasets. The AI is trained with one condition (or tag). Therefore, it is not possible to accurately indicate the fundus images of eyes with multiple diseases. The datasets are predominantly collected from East Asia. This may introduce a demographic bias. This can be resolved through a collection of data points from various datasets from different regions, but tele-optometry has not been extensively explored for there to be a large collection of publicly available high-quality datasets.

While this prototype can demonstrate the ability to determine eye-related diseases through AI, it will require a verification process by professionals to ensure that the results are accurate to be implemented as a medical tool. It is crucial for medical technologies that serve the general public to be reliable and not provide false or inconsistent test results because they can cause either unnecessary concerns or undetected critical issues.

Acknowledgments

I would like to thank Dr. Katherine Lai and Dr. Debora M. Lee Chen at the University of California, Berkeley School of Optometry and Claire N. Saunders at the California Institute of Technology for assisting me through the research process. Additionally, this submission would not be possible with Mr. Ryan Johnson at Portola High School.

References

American Optometric Association. (n.d.). See the Full Picture of Your Health with an Annual Comprehensive Eye Exam. Retrieved July 22, 2021, from https://www.aoa.org/healthy-eyes/caring-for-your-eyes/full-picture-of-eye-health?sso=y

Bivens, J., & Zipperer, B.. (2020, August 26). Health insurance and the COVID-19 shock: What we know so far about health insurance losses and what it means for policy. Economic Policy Institute. https://www.epi.org/publication/health-insurance-and-the-covid-19-shock/

Boyd, K. (2021a, June 28). What Is Diabetic Retinopathy? American Academy of Ophthalmology. https://www.aao.org/eye-health/diseases/what-is-diabetic-retinopathy

Boyd, K. (2021b, June 15). What Is Glaucoma? American Academy of Ophthalmology. https://www.aao.org/eye-health/diseases/what-is-glaucoma

Boyd, K. (2021c, July 7). What Are Cataracts? American Academy of Ophthalmology. https://www.aao.org/eye-health/diseases/what-are-cataracts

Bureau, U. C. (2020, September 15). Health Insurance Coverage in the United States: 2019. Retrieved from https://www.census.gov/library/publications/2020/demo/p60-271.html

Chen, Y. (n.d.). Retina dataset containing 1) normal 2) cataract 3) glaucoma 4) retina disease. GitHub. Retrieved July 23, 2021, https://github.com/yiweichen04/retina_dataset

Copeland, B. (2018, November 21). MYCIN. Encyclopedia Britannica. https://www.britannica.com/technology/MYCIN

Garvey, C. (2018, April). A framework for evaluating barriers to the democratization of artificial intelligence. In Thirty-Second AAAI Conference on Artificial Intelligence.

Guo, K., Ren, S., Bhuiyan, M. Z. A., Li, T., Liu, D., Liang, Z., & Chen, X. (2020). MDMaaS: Medical-Assisted Diagnosis Model as a Service With Artificial Intelligence and Trust. IEEE Transactions on Industrial Informatics, 16(3), 2102–2114. https://doi.org/10.1109/tii.2019.2937547

Hendrick, B. (2011, May 19). CDC: Many Americans Are Skipping Eye Care. WebMD. https://www.webmd.com/eye-health/news/20110519/cdc-many-americans-are-skipping-eye-care

Hulsey, A. (2020). A look into the impacts of tele-optometry expansionSharma. *Honors Theses.* https://scholar.utc.edu/honors-theses/338

Joint Shantou International Eye Centre. (2019, June 18). 1000 Fundus images with 39 categories. Kaggle. https://www.kaggle.com/linchundan/fundusimage1000

Microsoft Azure. (n.d.). Cognitive Services. Retrieved July 23, 2021, from http://azure.microsoft.com/en-us/services/cognitive-services/

Microsoft Azure. (2020, March 6). Select a domain for a Custom Vision project. https://docs.microsoft.com/en-us/azure/cognitive-services/custom-vision-service/select-domain

Mohajon, J. (2021, June 18). Confusion Matrix for Your Multi-Class Machine Learning Model. Medium. https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826

Mrutyunjaya, & Raga, S. (2020, October 14). A Smartphone Based Application for Early Detection of Diabetic Retinopathy Using Normal Eye Extraction. IEEE Xplore. https://ieeexplore.ieee.org/document/9277157

Noronha, K., & Nayak, K. P. (2012, February 1). Fundus image analysis for the detection of diabetic eye diseases-a review. IEEE Xplore. https://ieeexplore.ieee.org/document/6179013

Optix Family Eyecare. (2020, April 16). What Services Can I Get Using Tele-Optometry? https://www.optixeyecare.com/2020/04/16/what-services-can-i-get-using-tele-optometry/

Patry, G. G. G. (2021, February 18). Messidor-2. ADCIS. https://www.adcis.net/en/third-party/messidor2/

Rotterdam Ophthalmic Data Repository. (n.d.). Datasets. Retrieved July 23, 2021, from http://www.rodrep.com/data-sets.html

Primera Eye Care. (2020, November 1). 9 Things That Cause Bad Eyesight | Primera Eye Care. https://primeraeyecare.com/what-causes-bad-eyesight/

Shanggong Medical Technology Co., Ltd. (2020, September 24). Ocular Disease Recognition. Kaggle. https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k

Sharma, M., Jain, N., Ranganathan, S., Sharma, N., Honavar, S. G., Sharma, N., & Sachdev, M. S. (2020). Tele-ophthalmology: Need of the hour. Indian journal of ophthalmology, 68(7), 1328–1338. https://doi.org/10.4103/ijo.IJO_1784_20

Shortliffe E. H. (1977). Mycin: A Knowledge-Based Computer Program Applied to Infectious Diseases. Proceedings of the Annual Symposium on Computer Application in Medical Care, 66–69.

Taylor, H. R., Vu, H. T., McCarty, C. A., & Keeffe, J. E. (2004, August 1). The Need for Routine Eye Examinations. Investigative Ophthalmology & Visual Science. https://iovs.arvojournals.org/article.aspx?articleid=2124185

Turbert, D. (2021, April 9). Nearsightedness: What Is Myopia? American Academy of Ophthalmology. https://www.aao.org/eye-health/diseases/myopia-nearsightedness