OpthoAI:
A conversational multi-disease AI Ophthalmology smartphone screener with a novel approach to monitor disease progression

Alex Chirayath

Abstract

People who cannot afford annual eye checkups or do not have access to ophthalmologists are at high risk of preventable blindness from various undetected eye diseases, which presents the need for an easily accessible eye disease detector. OpthoAI is a low-cost Artificial Intelligence Ophthalmology screener that detects and monitors the progression of multiple common eye diseases, such as Drusens, Diabetic Retinopathy, Age-related Macular Degeneration, Hemorrhage, Glaucoma, Vascular Occlusion, Macular Edema, and Nevus. It utilizes an inexpensive lens apparatus to capture eye fundus images which can be uploaded into the conversational smartphone app. OpthoAI simulates an actual doctor-patient interaction, provides the patient with educational information, and recommends a follow-up appointment if needed. Chronic patients can reduce the frequency of follow-up visits to the ophthalmologist by using OpthoAI for interim follow-ups that check for treatment efficacy and disease progression. The dataset used for training the model was from the Brazilian Multilabel Ophthalmological Dataset, consisting of 16,266 Fundus images. Multiple disease-specific AI Convolutional Neural Networks (CNNs) were used for the initial disease diagnosis and progression, and each one used MobileNet as the base transfer learning model. Monitoring disease progression of a single disease solely using Artificial Intelligence is challenging and not as accurate as detecting different diseases. Therefore, this paper presents a novel approach to eye disease progression by combining Artificial Intelligence predictions based on eye images with additional inputs: an encoded Snellen eye chart and a patient's self-assessment/history. OpthoAI could successfully detect eight common eye diseases and measure the progression of Diabetic Retinopathy with a high confidence rate.

1. Introduction

Vision impairment or blindness can place a burden on an individual's quality of life and financial stability. According to the International Agency for the Prevention of Blindness, an estimated 1.1 billion people around the world live with vision loss. However, atleast 90% of vision loss is preventable or treatable. There is a vast inequality in the populations affected by vision loss, 90% of people with vision loss live in low and middle-income countries and 73% of people with vision loss are over 50 years old [1]. Therefore, this project developed 9 different transfer learning computer vision classification models, that deal with image data, to help prevent blindness in underserved areas with no Ophthalmologists or limited affordability, such as rural areas, developing countries, and point-of-care settings. The overall goal was to develop a comprehensive screener that can be employed on a mobile application and is effective from the initial diagnosis to lifelong monitoring of disease progression and treatment efficacy. Additionally, the Fundus images uploaded into the mobile application (OpthoAI) are provided by an inexpensive lens apparatus that was constructed in this project.

2. Background

As the prevalence of Artificial Intelligence grows, there has been an increase in the use of AI within the healthcare field, including ophthalmology. However, previous efforts in this field were limited to the detection of one disease at a time. On top of that, research doesn't go beyond the initial diagnosis of a disease. The proposed approach seen in this paper expands on these previous efforts by screening for multiple eye diseases at a time and implementing a novel approach for chronic/follow-up patients to monitor disease progression. A cheap and accessible way for eye disease detection and progression is a relevant issue globally because there is a general trend of there being a higher density of ophthalmologists in higher income areas compared to lower income areas [2]. Therefore, the AI disease-specific models and the progression algorithm will help prevent blindness by providing low-income individuals or individuals in remote areas with an accessible and low-cost ophthalmology disease screener that will catch eye diseases in their early stages. For the disease progression algorithm, the model is focused on the advancement of Diabetic Retinopathy and is combined with two human inputs, the Snellen Eye Chart test and a patient self-evaluation. Diabetic Retinopathy (DR) is an eye disease that affects the vision of diabetics by damaging the retina's blood vessels. An estimated ninety-three million people around the world have diabetic retinopathy, and it's the leading cause of vision loss globally [3]. Diabetic Retinopathy was used for progression due to its data availability, recognizable features, and five distinct ICDR classes. The ICDR (International Clinic Diabetic Retinopathy classification) classes are from a range of 0-4: 0 is no retinopathy, 1 is mild non-proliferative diabetic retinopathy, 2 is moderate non-proliferative diabetic retinopathy, 3 is severe non-proliferative diabetic retinopathy, and 4 is proliferative diabetic retinopathy and post-laser status.

3. Dataset

The dataset used for training and testing the models was the Brazilian Multilabel Ophthalmological Dataset (BRSET). The BRSET contains 16,266 Fundus images, which are images of the back of the eye including the retina, optic nerve head, macula, retinal blood vessels, choroid, and the vitreous [4]. Each image was 224 x 224. The dataset includes labels for each image including demographic information, structural label, diagnosis, and quality parameters labels, which can be accessed via the following link: https://doi.org/10.13026/xcxw-8198 [5].

4. Methodology/Models

4.1. Solution Overview

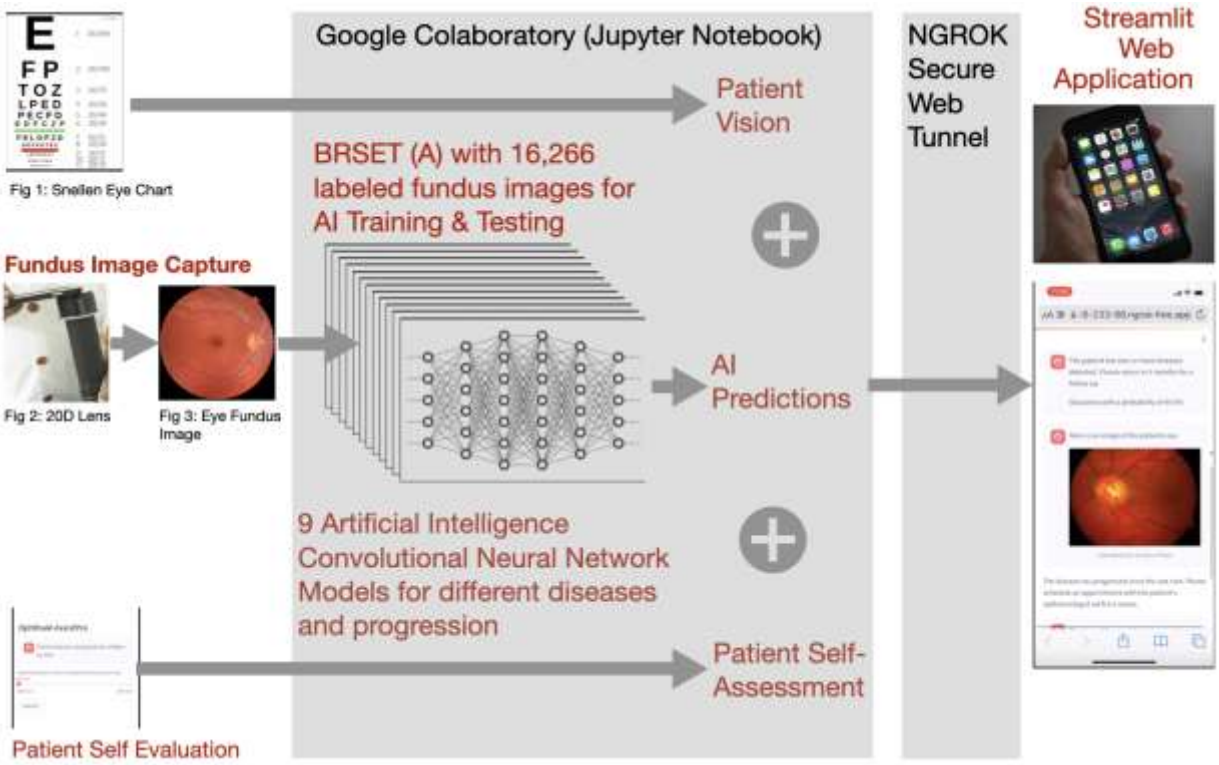

Figure 1: Project Overview

4.2. Imaging Prototype Device

The prototype of the device developed for capturing Fundus images can be made out of a 12-inch tube and a 20-diopter double aspheric lens. Inside the tube, black sheets of paper were placed to block out surrounding light. A piece of cardboard that fits the dimensions of an

average smartphone was cut out, and holes for the camera were also cut out. The pieces of cardboard were attached to a smartphone capable of capturing quality images and aligned to the hole cutout with the smartphone's camera. While in camera mode, the smartphone's flashlight was turned on to face the inside of the tube, and was held 2 inches over the subject's eye to capture the images. This lens apparatus device provides an inexpensive and non-invasive way to take Fundus images of a patient without the need for dilation drops.

4.3. Disease-specific Models and Progression Model Construction

Nine different Artificial Intelligence convolutional neural network models were developed, consisting of 8 models for the different diseases and one Diabetic Retinopathy model. Each model had a total of 32 layers; 28 of the layers were from MobileNetV1 followed by 3 fully connected dense layers and an output layer with a sigmoid or softmax activation function depending upon the type of classification. The disease-specific models used binary classification. The first step to creating the binary classification CNN model for each disease was to obtain and download the BRSET dataset. In Google Colaboratory notebooks, the necessary TensorFlow libraries were imported and the dataset was loaded in. The disease labels provided by the BRSET were then matched to each image. Downsampling and data synthesis were done if a disease did not meet a threshold of the number of samples in the minority class. For training and testing an 80/20 split was used. Mobilenet was used as a base and combined with custom layers to fit the dataset and the disease-specific labels.

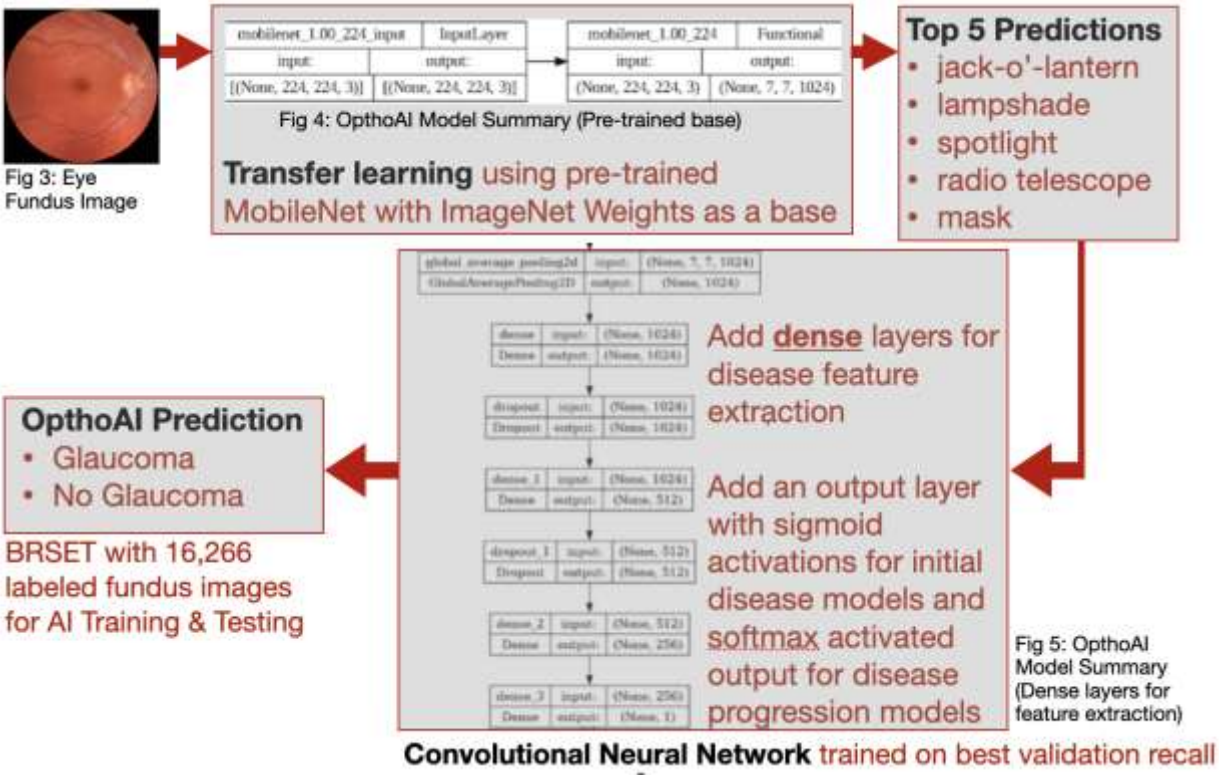

Figure 2: AI Model Architecture

4.4. AI Optimizations

To reduce bias preprocessing was done such as adding class weights, normalization, and data augmentation layers (RandomFlip, RandomRotation, RandomContrast, RandomZoom, RandomHeight, RandomWidth, RandomBrightness, RandomTranslation, and RandomShear). Dense and dropout layers were also added.

The batch size was set to 32, epochs were equal to 30, dropout was 25%, and the Sigmoid activation function was used. Early stopping was implemented, monitoring validation recall. After these parameters were set the model was compiled with Binary Cross Entropy as the loss function. The model was then fit to the training data and validated on the training data. A confusion matrix was used to evaluate how the model was performing for each class.

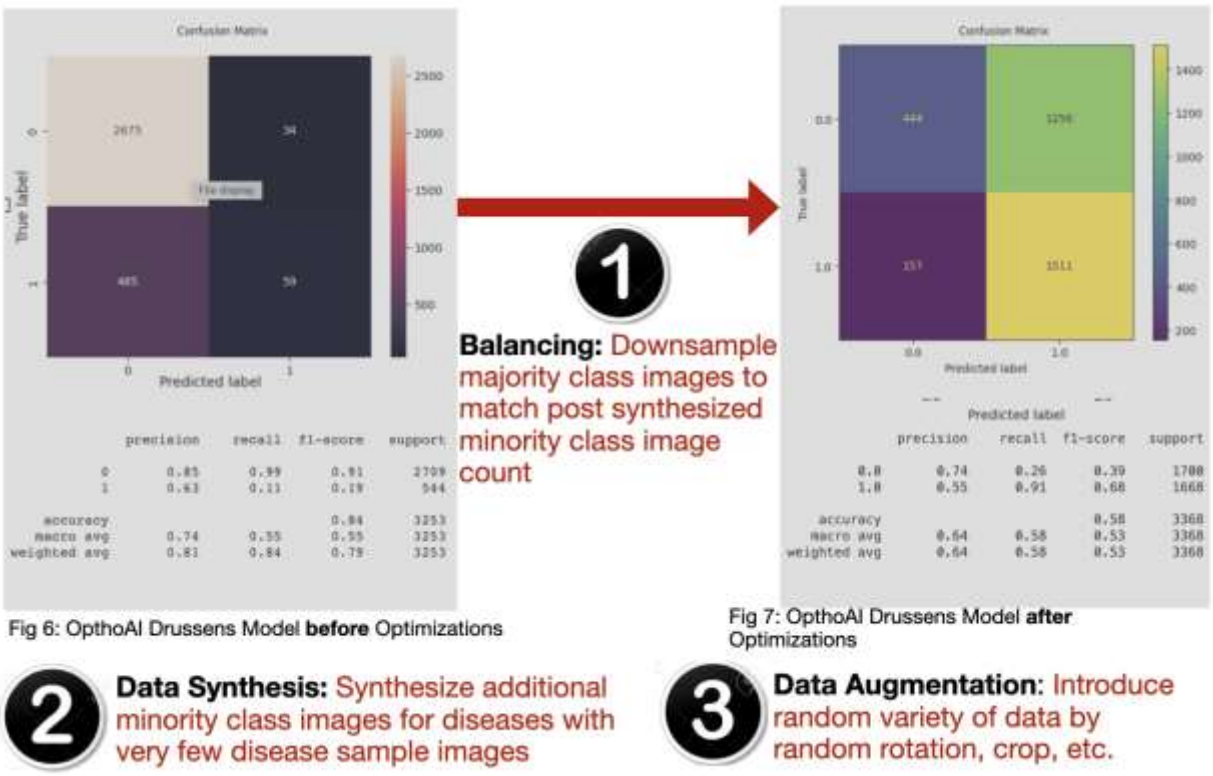

Fig 6: OpthoAI Drussens Model **before** Optimizations

Fig 7: OpthoAI Drussens Model **after** Optimizations

Figure 3: AI Optimizations

To create a model for a different disease (Glaucoma, Age-related Macular Degeneration, Macular Edema, Drusens, etc.), change the disease target to the label for the desired disease.

4.5. Progression

The progression model uses multiclass classification; there are 5 different classes (0 = no disease, 1 = mild non-proliferative disease, 2 = moderate non-proliferative disease, 3 = severe non-proliferative disease, and 4 = proliferative disease state). First, the dataset, labels, and necessary libraries were imported. The image labels were matched to each class (ID mapping). Following this, “DR_ICDR” was chosen as the disease target. Downsampling and data synthesis was done for classes 1-4 to closely match the 0 class (which had 2000 images). The progression model followed the same architecture as the disease-specific models except for the activation function being Softmax and the loss function being Categorical Cross Entropy.

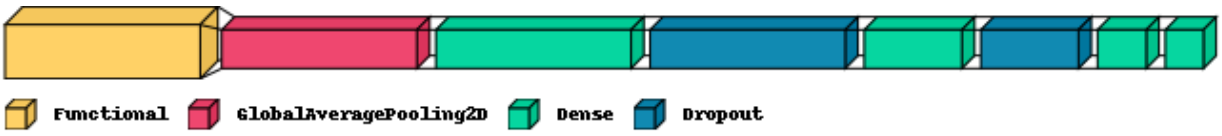

Figure 4: Diabetic Retinopathy Progression Model Architecture

4.6. Progression Algorithm and OpthoAI App Development

The OpthoAI mobile application was developed in a new Google Colaboratory notebook with the Streamlit and NGROK libraries imported. The novel approach to monitoring Diabetic Retinopathy progression was coded in the app’s notebook and includes 3 different ways to evaluate progression: the AI progression model, the Snellen Eye Chart test, and a patient self-evaluation. In this notebook, all 8 disease-specific models were uploaded together along with the DR progression model.

If the progression model predicts the patient to be in a higher ICDR class than they were in a previous visit then the app will give an output saying the disease has progressed overall. The app program goes through 5 different stages: Stage 0, Stage 5, Stage 10, Stage 20, and Stage 40. Stage 0 is the greeting and establishes if the patient is a new or returning patient. Stage 5 is the Snellen Eye Chart test for both new and returning patients. The best line of vision of the patient is recorded and put into a “points system”, 20/20 vision (perfect vision) is equal to 100 points, 20/40 is 80 points, 20/50 is 70 points, 20/70 is 50 points, 20/100 is 20 points, and 20/200 (legally blind) is 0 points. On top of that, for each letter the patient misses from their line of best vision, a point is subtracted from the total score. This Snellen score is encoded into a range of 0-100 and saved for comparison to the next visit. If there is a 20-point or more deterioration the app will give an output saying the disease has progressed. Stage 10 is for a returning patient coming back for a follow-up, in this stage the patient is asked to do the Snellen Eye Chart test

again and asked if their vision has become worse, better, or the same compared to the previous visit (patient self-evaluation). If the patient's self-evaluation reports a worse vision than before and a 15% probability of being in a higher ICDR class, then the app will give an output saying the disease has progressed. Stage 20 is where the Fundus image is uploaded (in the jpg or jpeg format) and the predictions/results are displayed. Stage 40 is to reset/restart back to stage 0. To make the user interface of the app easy to use, certain features were also included, such as conversation starters, a drop-down menu that displays a Snellen Eye Chart, sliders to input visual acuity, and a file uploader/displayer.

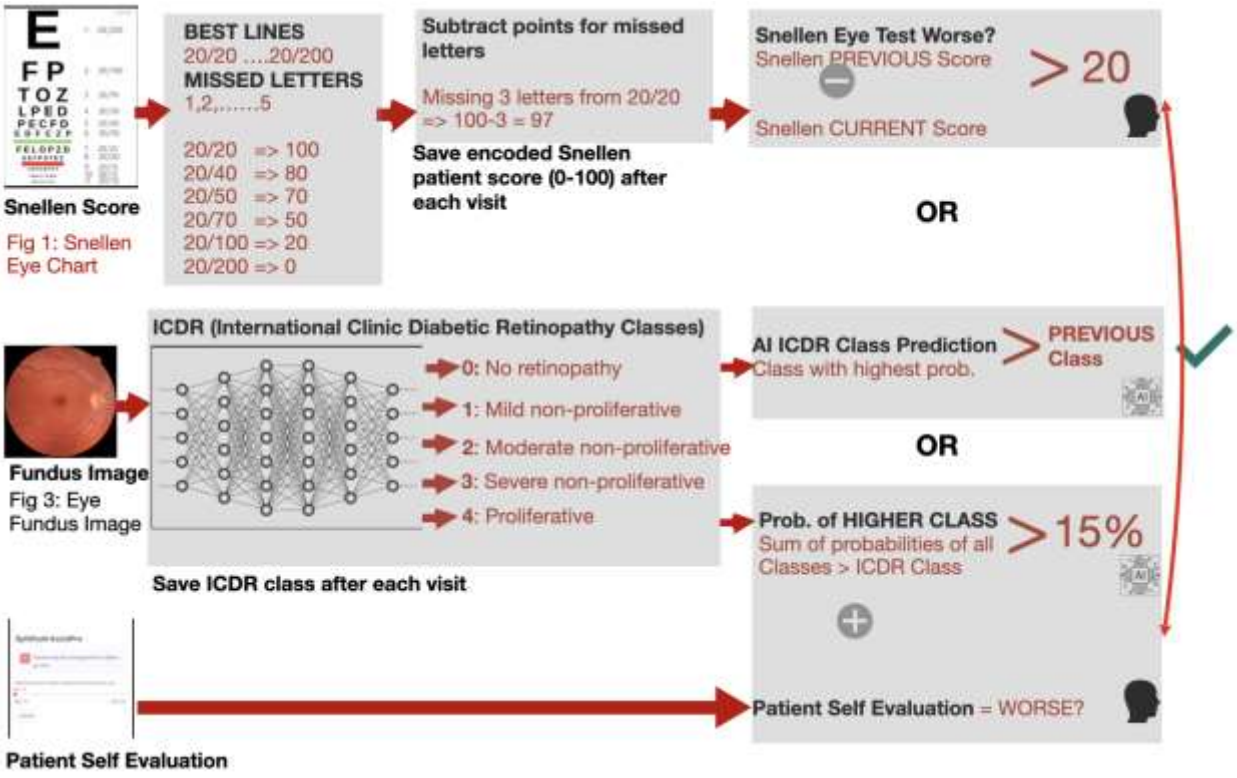

Figure 5: Disease Progression Algorithm

A demo of the OpthoAI user interface can be downloaded here: [OpthoAI Demo 02142024.mp4]

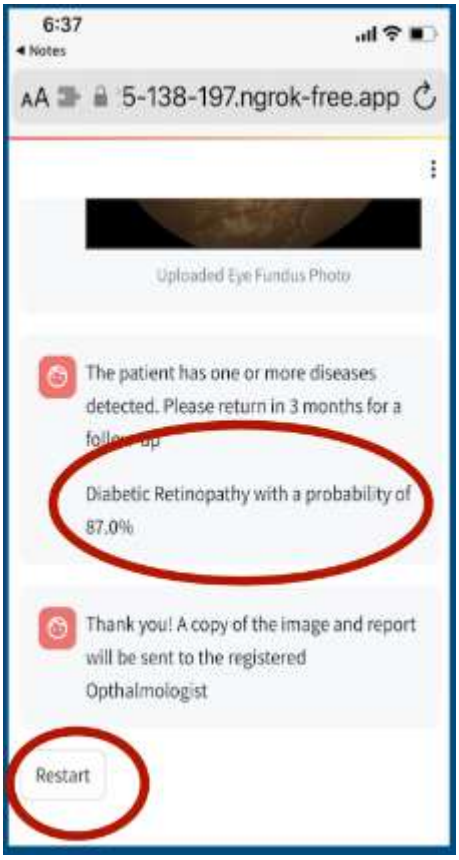

Figure 6: User interface of OpthoAI app

5. Results and Discussion

5.1. Rationale for Recall

Recall (true positive rate) was used to evaluate the model's performance because it minimizes false negatives and focuses on identifying the positive instances. Recall is a preferred metric for a screener in a real-life setting, because a false negative would mean that a patient had an eye disease that was not caught and the patient would walk away thinking their eyes were healthy. Conversely, a model with high recall may produce more false positives, but this is necessary for the need to mitigate false negatives, given their high cost in a healthcare scenario.

5.2. Disease-specific Model Results

Recall in the 1 class for Drusens is 91%. Recall in the 0 class is 26%. The model correctly predicted 1511 images out of 1668 images in the 1 class. Recall in the 1 class for Diabetic Retinopathy is 82%. Recall in the 0 class is 73%. The model correctly predicted 523 images out of 634 images in the 1 class. Recall in the 1 class for Age-related Macular Degeneration is 82%

Recall in the 0 class is 58%. The model correctly predicted 58 images out of 71 images in the 1 class. Recall in the 1 class for Hemorrhage is 66%. Recall in the 0 class is 83%. The model correctly predicted 58 images out of 88 images in the 1 class. Recall in the 1 class for Glaucoma is 71%. Recall in the 0 class is 75%. The model correctly predicted 436 images out of 617 images in the 1 class. Recall in the 1 class for Vascular Occlusion is 88%. Recall in the 0 class is 73%. The model correctly predicted 176 images out of 199 images in the 1 class. Recall in the 1 class for Macular Edema is 82%. Recall in the 0 class is 72%. The model correctly predicted 130 images out of 159 images in the 1 class. Recall in the 1 class for Nevus is 87%. Recall in the 0 class is 41%. The model correctly predicted 237 images out of 272 images in the 1 class.

	Recall 1 class	Recall 0 class
Drusens	91%	26%
DR	82%	73%
AMD	82%	58%
Hemorrhage	66%	83%
Glaucoma	71%	75%
Vascular Occ.	88%	73%
Macular Edema	82%	72%
Nevus	87%	41%

Table 1: Results for Disease-specific models

5.3. Progression Model Results

For the Diabetic Retinopathy progression model, the average recall is only 58%. For the 0 class the recall is 73%, class 1 is 35%, class 2 is 59%, class 3 is 68%, and class 4 is 55%. The model correctly predicted 213 images out of 407 images in the 0 class. The model correctly predicted 213 images out of 388 images in class 4. The model correctly predicted 242 images out of 355 images in class 3. The model correctly predicted 219 images out of 369 images in class 2. The model correctly predicted 122 images out of 351 images in class 1. The model correctly predicted 296 images out of 407 images in class 0.

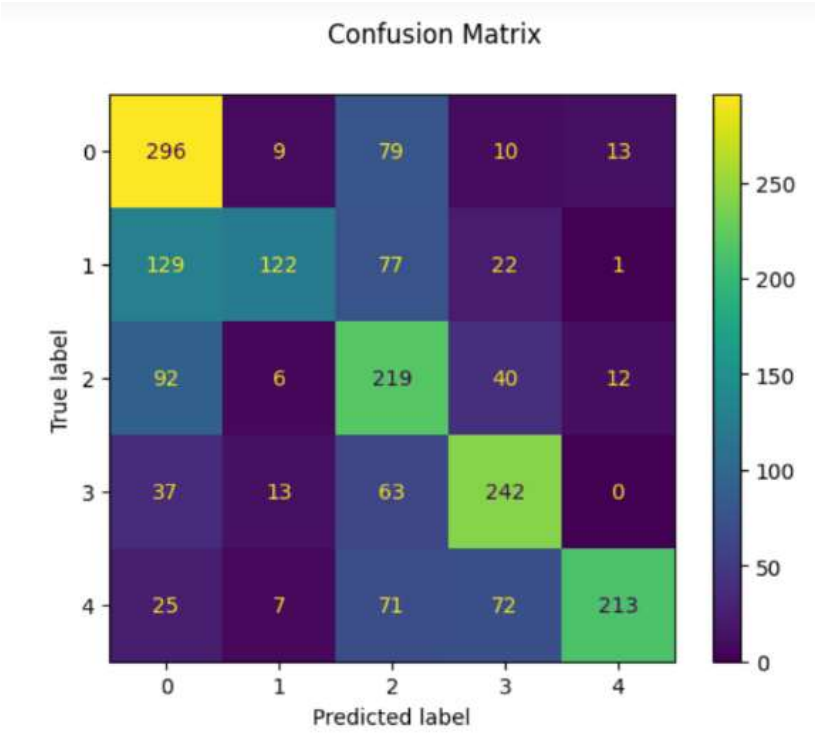

Figure 7: Confusion Matrix on Diabetic Retinopathy Progression Validation Dataset

5.4. Discussion

These disease-specific AI models show strong performance in detecting the presence of eye diseases from Fundus images. Overall the Drusens model had the highest performance in the 1 class, this is likely due to the high amount of images it had for each class (1668 in the 1 class and 1700 in the 0 class). This underscores the importance of having a sufficiently large and representative dataset for training robust diagnostic models, as models that were trained on fewer and more imbalanced data performed poorer.

The progression model for Diabetic Retinopathy had the lowest recall at 58% and the poorest performance across the 5 ICDR classes, notably the recall for class 1 (mild non-proliferative DR) is 35%, which suggests the model's struggles to detect the early signs of the disease. However, performance improves for more severe stages, peaking at 68% for class 3 (severe non-proliferative DR), indicating that the model may be useful for flagging advanced cases for referral, but is less reliable for catching the disease early.

To compensate for the limitations of the stand-alone progression model, this project took the novel approach of combining AI predictions with two human-generated inputs of visual acuity from a Snellen Eye Chart and a self-assessment of vision changes. Deterioration of 20 or more points of the Snellen score

compared to the previous visit indicates the progression of the disease. If the model predicts the patient has at least a 15% chance of an increase in their severity class and a worse self-assessment, it's reported as having disease progression. This multi-pronged approach aims to capitalize on the pattern recognition capabilities of AI and the patient's self-recorded evaluation to enable more sensitive monitoring of chronic Diabetic Retinopathy. Further clinical validation is needed to further assess the real-world effectiveness of this hybrid algorithm.

Some key limitations of this study include the narrow geographic scope of the dataset, as all images come from Brazilian patients as well as the app being in the prototyping stages of development. Priorities for future development of the app include: expanding the training data to enhance global generalizability, refining the app's user interface and performance, providing localized patient education/resources, and continuing to optimize the model architecture and the progression algorithm. Lastly, with further advancement, OpthoAI has the potential to reduce preventable blindness by extending access to eye disease screening and monitoring, especially in low-resource settings with limited availability of eye-care specialists.

6. Conclusion

This project was able to develop Artificial Intelligence models that successfully detects 8 prevalent eye diseases with a total average recall of 83% in the 1 class on the validation datasets. In addition, disease progression for Diabetic Retinopathy was proven possible by augmenting a 58% recall AI model with a novel algorithm that combines AI predictions with human inputs from an encoded Snellen Eye Chart test and patient self-evaluation. An intuitive and and conversational web application, called OpthoAI, was also developed. This mobile app has an easy-to-understand user interface and requires limited training. To provide an inexpensive way to capture Fundus images a prototype lens apparatus was made using a 20 Diopter lens. To expand/improve on this research I would like to include different demographics for the dataset, such as The Indian Diabetic Retinopathy Dataset [6]. The BRSET used to train the models was limited to only Brazilian patients. I also aim to enhance the performance of the OpthoAI app, such as faster runtime and more features. In the app, I also intend to provide patient education if a disease is predicted. For the models, I aim to improve the metrics (precision, recall, or accuracy) and minimize false positives. Additionally, I want to implement a weighted average for the ICDR progression score instead of the discrete values used currently.

7. Data Availability

The code for the project as well as an OpthoAI demo video can be accessed at the following link: https://github.com/Alxchirayath/OpthoAI/tree/main.
If there is an error with accessing the notebooks with the previous link, then it can also be accessed here: https://nbviewer.org/github/Alxchirayath/OpthoAI/tree/main/

References

[1]
The International Agency for the Prevention of Blindness. Vision Atlas - the International Agency for the Prevention of Blindness. The International Agency for the Prevention of Blindness. Published December 18, 2023. https://www.iapb.org/learn/vision-atlas/

[2]
Resnikoff S, Lansingh VC, Washburn L, et al. Estimated number of ophthalmologists worldwide (International Council of Ophthalmology update): will we meet the needs? British Journal of Ophthalmology. 2019;104(4):588-592. https://bjo.bmj.com/content/104/4/588

[3]
Shukla UV, Tripathy K. Diabetic retinopathy. StatPearls - NCBI Bookshelf. Published August 25, 2023. https://www.ncbi.nlm.nih.gov/books/NBK560805/

[4]
Mishra C, Tripathy K. Fundus Camera. StatPearls - NCBI Bookshelf. Published August 25, 2023. https://www.ncbi.nlm.nih.gov/books/NBK585111/

[5]
Nakayama LF, Restrepo D, Matos J, et al. BRSET: a Brazilian multilabel ophthalmological dataset of retina Fundus photos. Published 2024. https://www.semanticscholar.org/paper/BRSET%3A-A-Brazilian-Multilabel-Ophthalmological-of-Nakayama-Restrepo/2fae0d68ba7aa30db59717ab2aed6ace554ba8ef

[6]
Porwal P, Pachade S, Kamble R, et al. Indian Diabetic Retinopathy Image Dataset (IDRID): a database for diabetic retinopathy screening research. Published July 10, 2018. https://doi.org/10.3390/data3030025