SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Ocular Disease Recognition

Design Track

50.035: Computer Vision Final Report

7th December 2023

Team 10

Shawn Choo	1005128
Chen Jiasen	1005220
Yeo Wan Li	1005321
Claudia Lai	1005311
Samuel Chiang	1005142
Gabriel Teng	1005027

Table of Contents

Section 1: Introduction

1.1 Background

In today's society, good health and well-being are some of the few platforms on which nations can come together to achieve a shared objective. The epitome of this commitment on a global stage is the 2030 Agenda for Sustainable Development (GENeco., n.d.), where UN member nations identified Good Health and Well-being as one of the 17 Sustainable Development Goals (SDG) to ensure healthy lives and promote well-being for all at all ages (UN, 2015). As Singapore is an ageing society, ocular diseases such as diabetic retinopathy (Diabetic Retinopathy in Singaporean population, 2023), glaucoma (Glaucoma in Singapore: Stats, Risk Factors and Prevention, n.d.), and age-related macular degeneration (Age-related Macular Degeneration, n.d.) are becoming more commonplace (Relationship between vision impairment and employment, n.d.). This acceleration in poor eye health underscores the need for early intervention to prevent further visual impairment.

This is especially important in Singapore whereby its workforce is the country's primary natural resource (The secret of Singapore's success in education, 2015). In fact, from a study conducted by the Singapore National Eye Center, it was found that for Singaporeans aged ≥ 40 years with mild vision loss had an approximately 50% chance of being unemployed at baseline (Relationship between vision impairment and employment, n.d.).

However with the high cost of an eye appointment in Singapore; starting from a base screening cost of $50 and above (Eye Screening Singapore, n.d.) and the long and tedious booking procedures, many people choose not to book eye appointments due to these barriers to entry.

Therefore, for us to ensure a healthy working population, the population's good wellbeing and health should be prioritized to maintain workforce efficiency and support economic growth. As such, ocular health plays a key role in ensuring the employability of our workforce especially when many people in Singapore use the computer.

Computer vision plays a pivotal role in this project as it forms the foundation for possible symptoms of ocular diseases through visual data. The combination of deep learning models and computer vision enables us to bridge the gap between raw visual data and actionable information.

1.2 Data

We managed to source a dataset from Kaggle that is applicable for our use cases. We managed to find multiple Retinal Disease Classification images:

1. Retinal Disease Classification (3200 fundus images for 46 different diseases)
2. Ocular Disease Recognition (Right and left eye fundus photographs of 5000 patients)
3. Eye_diseases_classification (Eye Disease Retinal Images)

Eventually, we decided to proceed with using the eye_diseases_classification due to the size of the dataset which directly affects our training progress. Retinal Disease Classification is 8 Gb in size, Ocular Disease Recognition is 2Gb in size and eye_diseases_classification is 771 Mb in size. Thus by considering the time frame as well as computing power, we have chosen the eye_diseases_classification dataset.

The dataset consists of Normal, Diabetic Retinopathy, Cataract and Glaucoma retinal images (see Figures A1 to A4, Appendix A), where each class has approximately 1000 images. These images are collected from various sources like IDRiD, Ocular recognition, HRF etc (Kaggle, 2022), which are more suitable for our use case as these are the most common eye diseases in Singapore.

Section 2: Methodology

2.1 CNN

Convolutional Neural Network (CNN) is a neural network which is designed to take images as in inputs and assign labels to them. Having multiple convolution layers allows the model to learn and understand, through feature extraction, the important features relevant in classifying and identifying eye disease type and presence.

2.1.1 Batch Normalization

In deep neural networks, the parameters from the previous layers may change after each mini-batch when the weights are updated. As such, the distribution of each layer's inputs also changes during training, which is known as covariate shift. This significantly slows down the learning rate as the network cannot generalize and be efficient. Batch normalization is used after each convolutional layer to address this issue by normalizing input data for each mini-batch (Brownlee, 2019). This technique stabilizes the learning process, reducing the number of training epochs required to train the network.

2.1.2 Dropout

Dropout is an effective approach to address overfitting that may arise in training deep neural networks. Dropouts are added to randomly switch a specified percentage of neurons in the network, thereby switching off incoming and outgoing connections to those neurons (Dwivedi, 2022). While the number of iterations needed will increase, training with dropout enhances the overall learning of the model as the speed for running each epoch in training also increases.

2.1.3 Activation Functions

Activation functions determine how the weighted sum of the inputs is transformed into an output node in a layer of the network (Brownlee, 2021). This is to introduce non-linearity to the output of neurons. Using activation functions make back-propagation possible as gradients are supplied to update the weights and biases. In our model, 3 activation functions are used. ReLU is used for convolutional layers, Softmax for the classifier layer and Sigmoid in the ‘by-disease’ classification.

2.1.4 Max Pooling

Max pooling layers reduces the dimensionality of input images by reducing the number of parameters. When the filter kernel moves across the input image, the pixel with the maximum value is selected for each iteration. Max pooling is helpful to reduce overfitting of images. This is because only the most activated pixels are selected and these high values are preserved while the lower valued pixels are discarded.

2.1.5 Adam Optimizer

Adam Optimizer is an iterative optimization algorithm used to minimize the loss function during training (Vishwakarma, 2023). Adam Optimizer adaptively adjusts the learning rate for each parameter and implements bias correction to counter initialization bias. This allows for faster convergence to optimal solutions and makes the model more invariant to diagonal rescaling of gradients. Hence, larger data parameters can be used in the model.

2.2 Flask Framework

As our project directive was to provide an easily accessible model for users to interact with the model, a Flask web application was developed to provide a proof of concept. Web applications can be easily viewed from any browser (mobile or otherwise) and thus does not need to be installed, hence easing access for various types of users.

2.2.1 Flask Pages

The Flask app consists of 2 main pages, the upload page and the result prediction page. The Flask app imports a model file (.h5 format) that was generated from training, and uses it to predict conditions based on a user-inputed image.

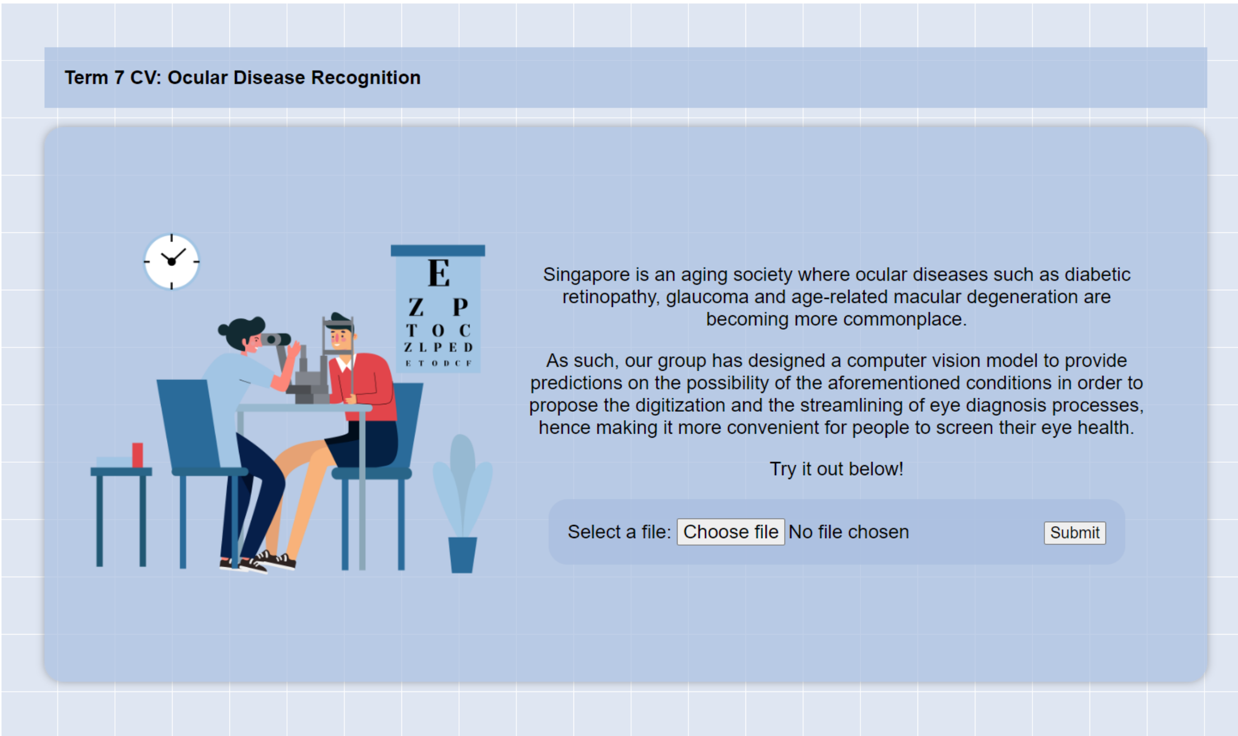

Figure 1: User-friendly interface for ease of uploading of retinal image

The user first starts by pressing on the "Choose file" button to upload their image. This opens up the file explorer of the user's device for them to choose their image to submit. Accepted file types include .jpg, .jpeg and .png.

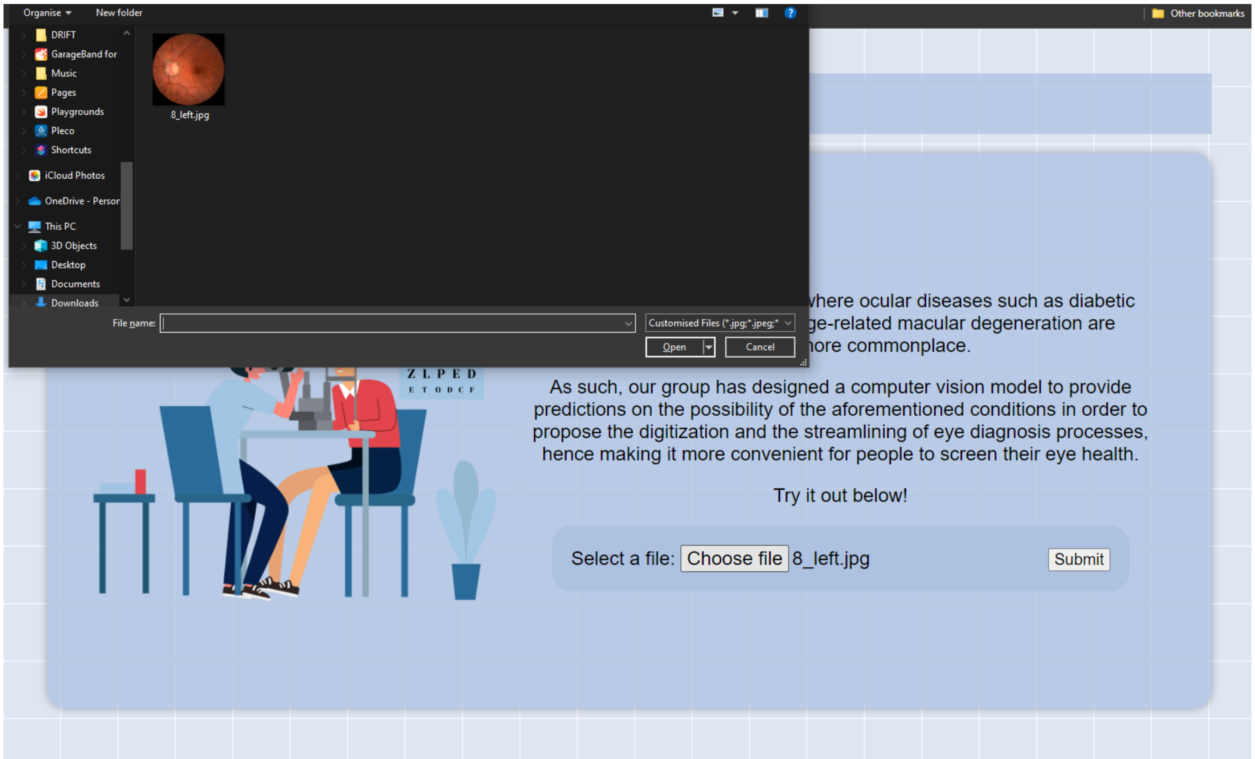

Figure 2: Uploading of our test image, 8_left.jpg

After choosing and uploading their image, the user then clicks on the "Submit" button, after which the Flask app will receive the file and use the imported model to generate a prediction on it.

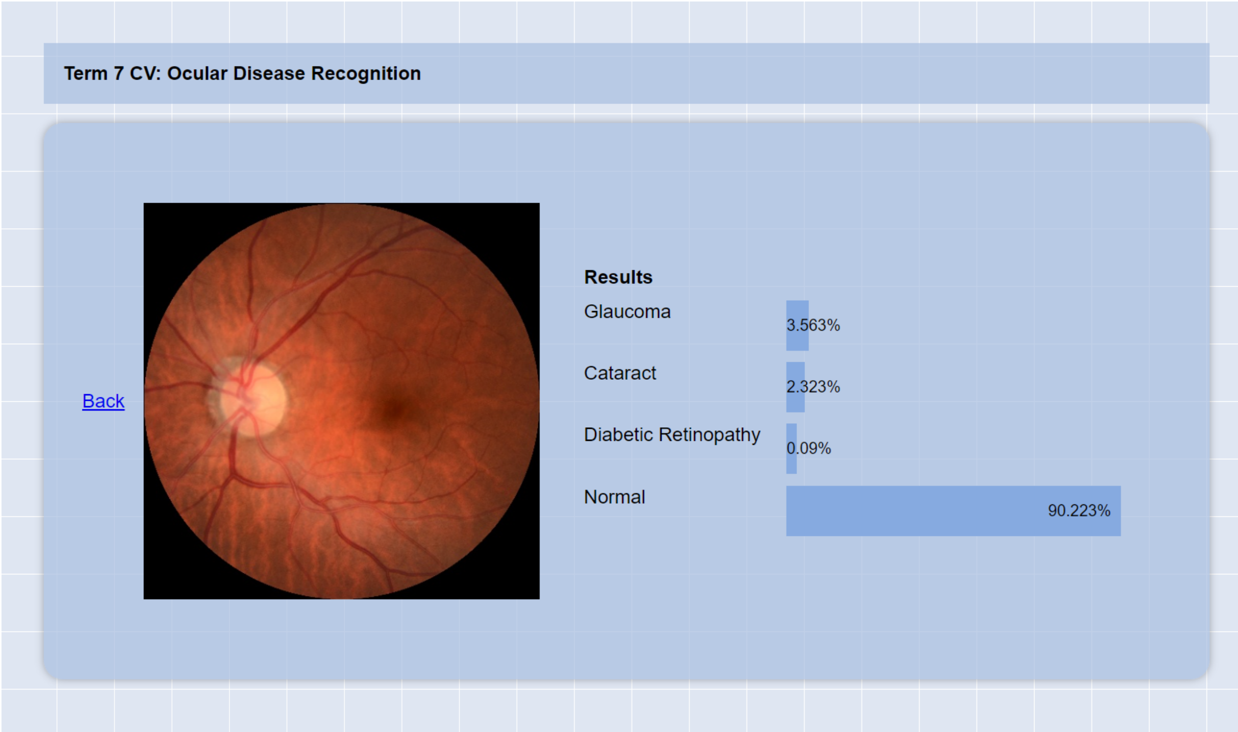

Figure 3: Results of model prediction displayed clearly for the user

As seen in the final page of this prototype, the probabilities of occurrence of the various conditions are displayed in terms of percentage. As our project did not involve professional medical input, we did not display concrete diagnoses of the uploaded image on the basis that our model has not been cleared to do so.

Section 3: Novelty Assessment of Solution

As it is uncommon for people to see an eye specialist unless signs or symptoms of eye diseases have already started showing, detecting and diagnosing such conditions would often be delayed due to inconvenience and a lack of awareness from the affected person. As such, the solution our group has would enable such convenience, as the accessibility of our solution is very open. Being a light-weighted model in size, the classifier model would be able to run on mobile devices or on the Flask web app, allowing people to casually check their eyes to see if there are any risks at their own convenience.

Currently, the latest solution for simplifying and easing the eye checkup process would be self-help kiosks, which are planted in specific eye specialist clinics or hospitals, which would free up doctors and specialists time, allowing them to diagnose and check on more patients. However, on the patient's end, they still have to make their way for such a self check up, and would only do so when they think they might have an eye issue, which may be late.

As such, our solution, which can be run on a mobile device, allowing patients to take their own photo of their eyes with the help of an attachable magnifier to their camera lens and letting the app with the classifier model determine the risk percentages as feedback for the user, would be a convenient and easy method for users to simply check their eyes.

Section 4: Work Assignment

Literature Review	Everyone
Ideation Fine Tuning	Everyone
UIUX Framework	Yeo Wan Li
CNN Architecture and Code Cleaning	Shawn Choo, Gabriel Teng, Claudia Lai
Flask Framework	Yeo Wan Li
Report Writing	Everyone
Slides	Everyone

Table 1: Work Assignment

The team's work assignment is based on tasks. A brief summary of the tasks allocated is shown in Table 1 above.

Section 5: Implementation and Results

5.1 Results of CNN

For our first prototype, we used Softmax to be our classifier layer. When given an input image, this allows the model to produce an output label from the list of the four target ocular diseases. However, the team later decided that the model should display probabilities of the four diseases to the user, instead of merely providing a predicted label of the disease. As such, we omitted the Softmax classifier layer and made modifications for the model to produce four fully-connected layers that would be concatenated as the final output.

The two CNN model architectures are shown in Figures A5 and A6, Appendix A. A summarized form of the models' formula is shown below when running $model.summary()$:

1. Conv2D-Pool-Conv2D-Pool-Conv2D-Pool-Conv2D-Pool-FC-FC-FC-Softmax
2. Conv2D-Pool-Conv2D-Pool-Conv2D-Pool-Conv2D-Pool-FC-FC-FC(x4)-Sigmoid(x4)-Concat

The ReLU activation function is used for all convolutional layers and fully-connected layers and batch normalization is used for each of these layers. This is followed by using dropouts with values of 0.25 for each batch normalization. Adam Optimizer is applied to the models as well, with a learning rate of 0.00025. To allow the modified model to classify multiple classes, Sigmoid activation function is used instead of Softmax.

Accuracy and Loss were plotted against Epoch for both models (see Figure A7 to A8, Appendix A).
For the Multiclass Classification (modified) model, some of these convolutions give very good feature extractions, and we trained the new 'by-disease' classifiers which use these features to give us the score that we want, 87% was what we got after training the model after 300 epochs and running the model twice. First run we obtained 83% and eventually for the second it increased to 87%. Both runs used the same test size from the dataset.

5.2 Classification Report

	Precision	Recall	F1-score	Support
Glaucoma	0.83	0.78	0.80	196
Cataract	0.90	0.99	0.94	206
Normal	0.73	0.83	0.78	223
Diabetic_retinopathy	0.88	0.71	0.79	218
Micro avg	0.83	0.83	**0.83**	843
Macro avg	0.83	0.83	0.83	843
Weighted avg	0.83	0.83	0.83	843
Samples avg	0.83	0.83	0.83	843

Table 2: Classification Report of CNN model's performance on Test Set

	Precision	Recall	F1-score	Support
Glaucoma	0.86	0.58	0.69	211
Cataract	0.77	0.90	0.83	199
Normal	0.93	1.00	0.97	228
Diabetic_retinopathy	0.76	0.83	0.80	206
Micro avg	0.83	0.83	**0.83**	844
Macro avg	0.83	0.83	0.82	844
Weighted avg	0.83	0.83	0.82	844
Samples avg	0.83	0.83	0.83	844

Table 3: Classification Report of By-Disease Predictor's performance on Test Set

The classification report in Table 2 above shows a representation of the main classification metrics (precision, recall, and F1 score) for each of the four class labels. The F1 score provides a combined view of both precision and recall metrics and reaches its maximum value when precision is equal to recall. As such, it is useful in cases where false positives and false negatives are equally costly (Agrawal, 2023). From the classification report, it is observed that the model reaches 83% for its F1 score.

5.3 Confusion Matrix

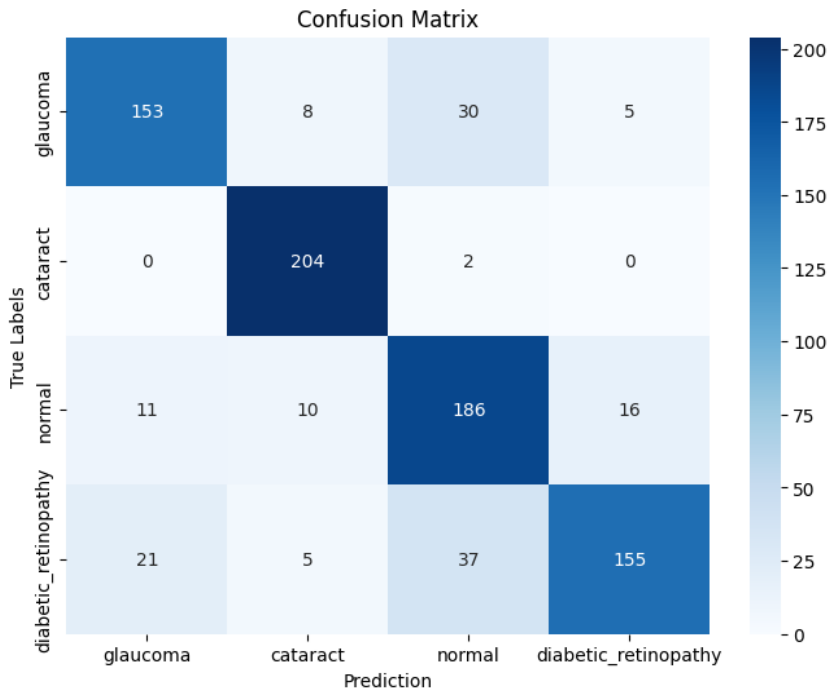

Figure 5: Confusion Matrix of CNN model's performance on Test Set

The confusion matrix, shown in Figure 5, is also obtained for a better visualization of the CNN model's performance from the testing results. From the matrix, it is observed that the model is able to predict the four classes accurately, of which it performs the best at detecting the cataract class.

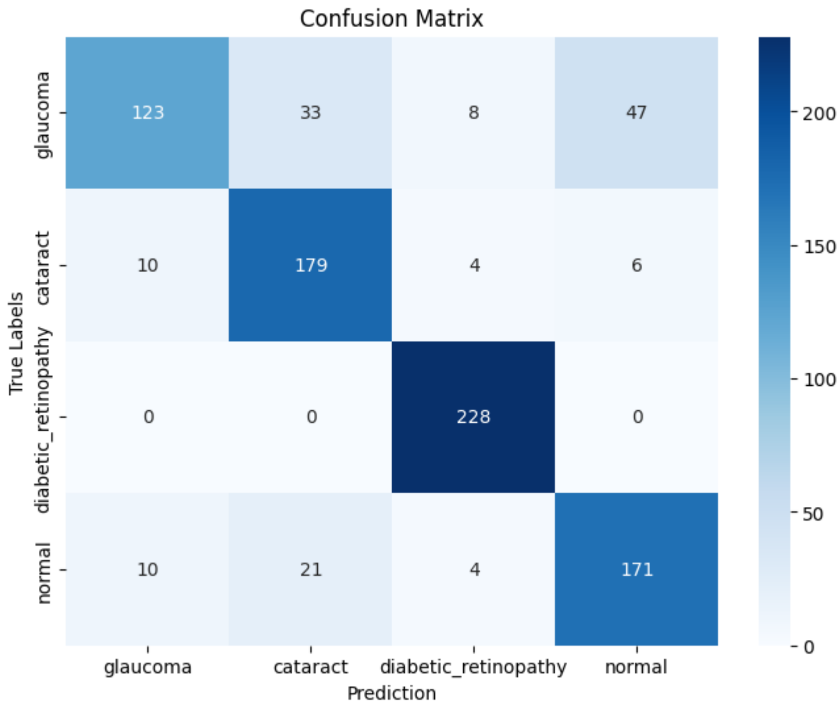

Figure 6: Confusion Matrix of By-Disease Predictor's performance on Test Set

In figure 6, it is observed that the By-Disease Predictor is able to predict the four classes accurately, of which it performs the best at detecting the diabetic retinopathy class.

5.4 Other Models Tested

5.4.1 DeeperConv

We added an extra convolution layer and an extra fully connected layer, however the accuracy for the DeeperConv model dropped to 62%.

5.4.1 ShortConv

We tried to reduce the depth of the model but increase the number of filters per convolution layer. The obtained accuracy of 81% is still lower than our By-Disease predictor.

Section 6: Conclusion and Future Work

OUR MODEL

Eye disease detection can be challenging, especially when trying to classify eye images of early stages, as there is a minute difference and a lack of such data as compared to a normal eye image. However, through our discoveries, we are pleased that the model is able to perform reasonably well, with at least an average F1-score of 83% for each of the 4 categories. And since it is better to have false positives than false negatives in the case of disease detection, from the test results we can see that glaucoma and diabetic retinopathy detection would need improvement in this aspect, while cataract detection is performing well. To improve on this, the threshold line for glaucoma and diabetic retinopathy activation can be lowered below the basic 0.5 in order to reduce the likelihood of false negatives. Having defined weights for each class can also alter the loss function and force the model to learn more about the class with heavier weights, preventing false negatives.

Overall, more techniques and architectures can be explored and implemented for future work. Such as the Resnet architecture type, which utilizes skip-connections to prevent vanishing gradients during training, or making use of generative adversarial networks in order to create more images for the training dataset. Images of eyes with conditions in early stages would also be valuable for training the model in order to predict and classify early cases, which is always better for disease detection.

OUR DASHBOARD

All in all, running our model on the dashboard would allow the general public to easily take a photo of their eyes with the help of an attachable magnifier to their camera lens, and all this from the comforts of their homes. As such, the convenience and accessibility of our solution overcomes barriers to entry by early detection of any potential ocular diseases and allows our classifier model to determine the risk percentages as feedback for users. Albeit difficult to overcome the natural human nature in diverting changes, our app aims to be the initial confirmation and detection of the potential of ocular diseases, and hence the worry for one’s health would give him or her the necessary impetus in going for an eye checkup. As such, we are confident that this would encourage users to overcome their initial inertia in taking the necessary step in getting a proper eye checkup at the clinic.

Given more time, we would test out the app with the general public and use success metrics such Neilson’s 10 usability heuristics for good interface design as to better understand how to improve our frontend to better achieve our goal in maintaining the health of our country’s workforce (Nielsen, 2020).

Appendix A

Figure A1: Glaucoma dataset images

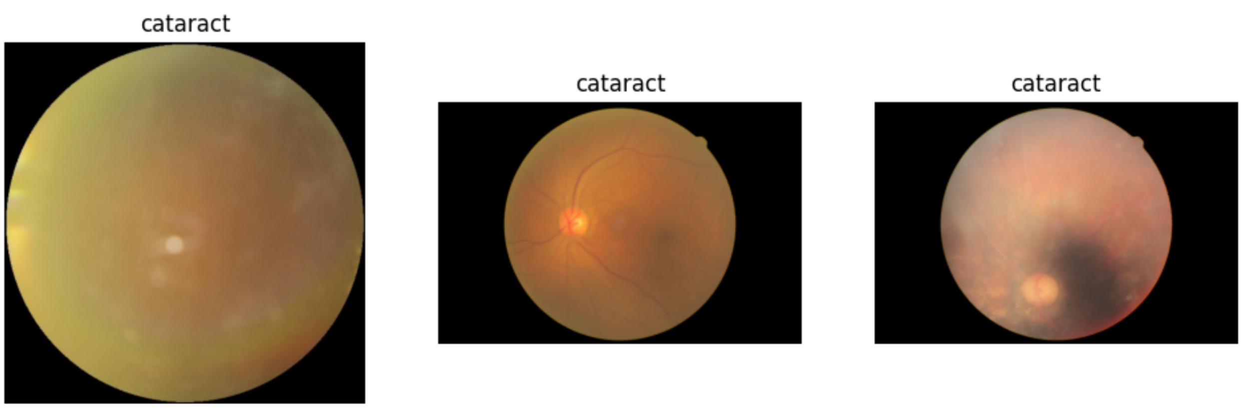

Figure A2: Cataract dataset images

Figure A3: Normal dataset images

Figure A4: Diabetic retinopathy dataset images

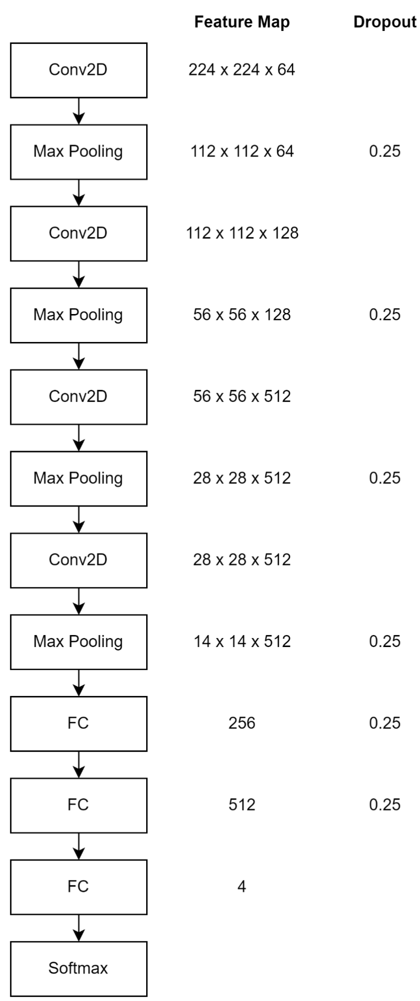

Figure A5: Initial CNN Model Summary

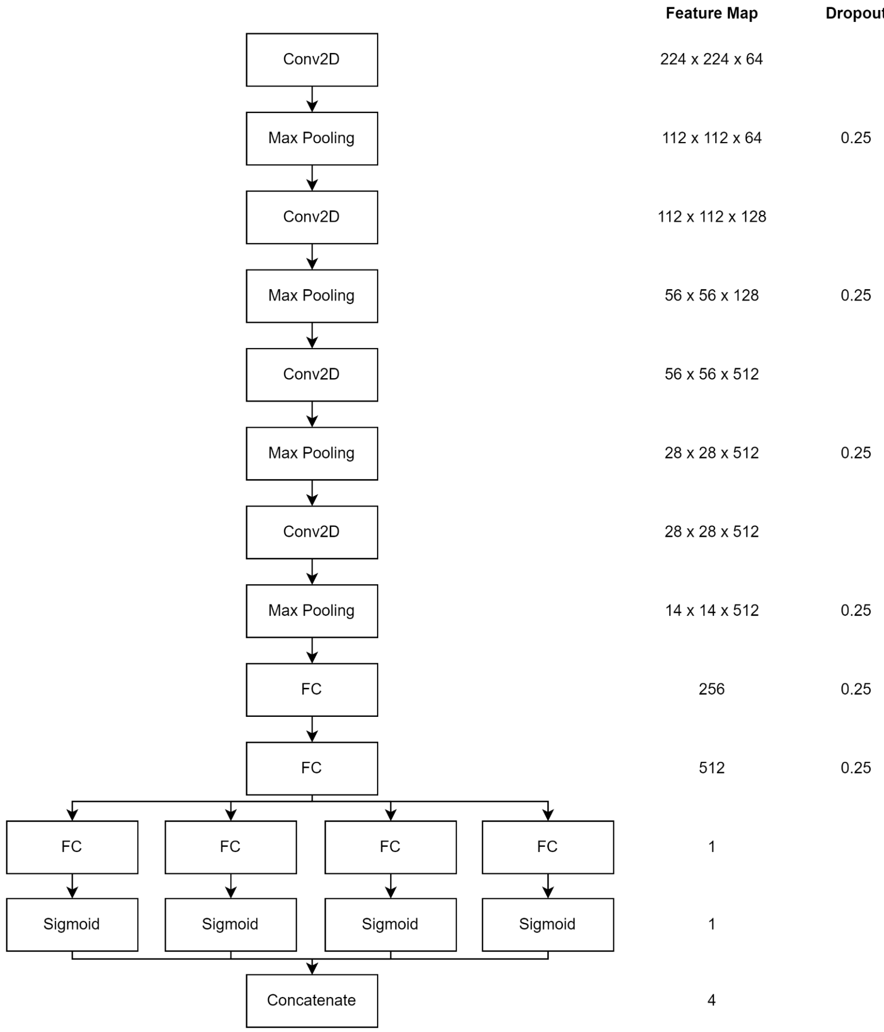

Figure A6: New CNN Model Summary

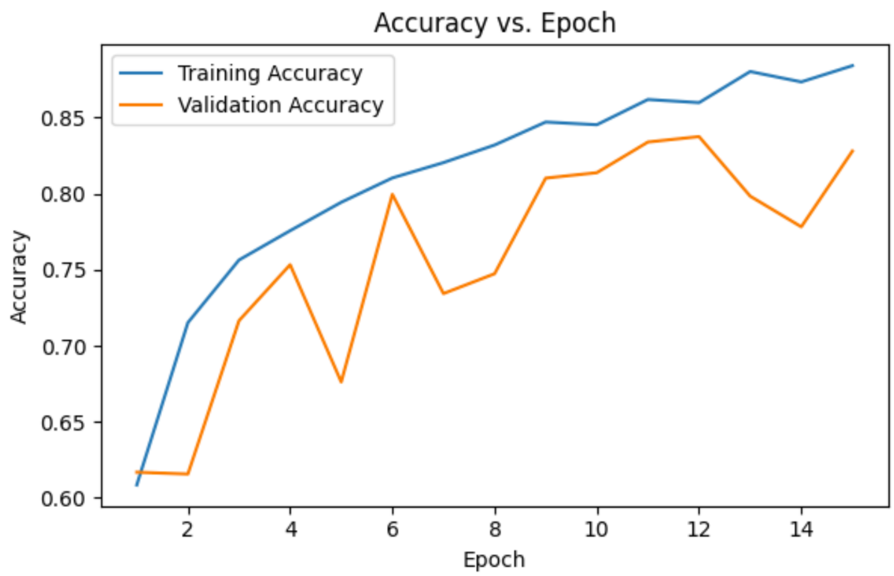

Figure A7: Graph of Accuracy against Epoch

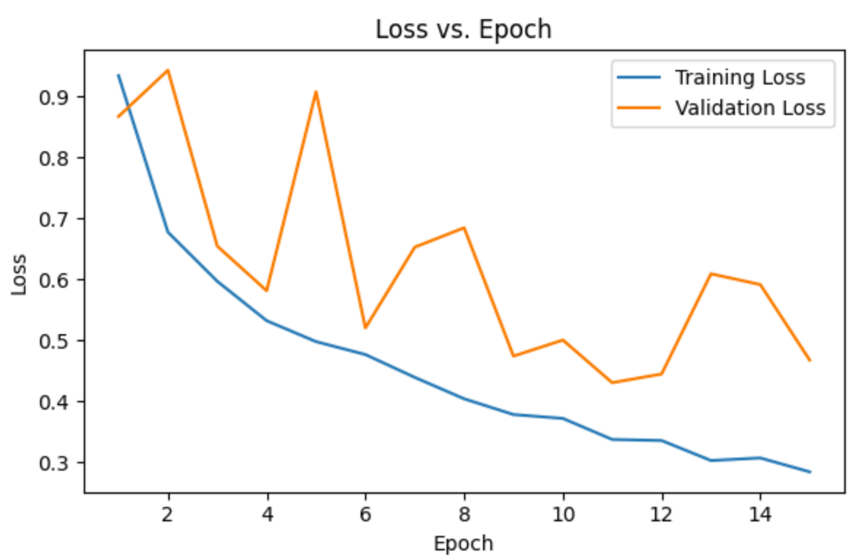

Figure A8: Graph of Loss against Epoch

Appendix B

Github link: https://github.com/jiasenn/TLDR

References

Diabetic Retinopathy in Singaporean population | OPTH. Dove Medical Press. (2023, February 2). Retrieved September 27, 2023, from https://www.dovepress.com/association-of-triglyceride-glucose-index-with-prevalence-and-incidenc-peer-reviewed-fulltext-article-OPTH

Glaucoma in Singapore: Stats, Risk Factors and Prevention. HealthXchange.sg. (n.d.). Retrieved September 27, 2023, from https://www.healthxchange.sg/seniors/ageing-concerns/glaucoma-singapore-stats-risk-factors-prevention

Age-related Macular Degeneration. Ng Teng Fong General Hospital. (n.d.). Retrieved September 27, 2023, from https://www.ntfgh.com.sg/Health-Information/Documents/brochures/Age-related%20Macular%20Degeneration.pdf

Research: Visual impairment and employment. EyeSight Issue 1 2022 - Relationship between vision impairment and employment. (n.d.). https://www.snec.com.sg/eyesight-issue1-2022-relationship-between-vision-impairment-and-employment

The secret of Singapore's success in education. Straits Times. (2015. April 11). Retrieved December 14, 2023 from https://www.straitstimes.com/opinion/the-secret-of-singapores-success-in-education

Eye Screening Singapore. HealthScreening.sg. (n.d.). Retrieved September 27, 2023, from https://healthscreening.sg/eye-screening-singapore

Larxel. (2021, August). Retinal Disease Classification, Version 1. Retrieved December September 27, 2023 from https://www.kaggle.com/datasets/andrewmvd/retinal-disease-classification

Larxel. (2020, September). Ocular Disease Recognition, Version 2. Retrieved December September 27, 2023 from https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k

Guna, V. D. (2022, August). eye_diseases_classification, Version 1. Retrieved December September 27, 2023 from https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification

Brownlee, J. (2019, December 3). A gentle introduction to batch normalization for Deep Neural Networks. MachineLearningMastery.com. https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/

Dwivedi, R. (2022, July 16). Everything you should know about dropouts and batchnormalization in CNN. Analytics India Magazine. https://analyticsindiamag.com/everything-you-should-know-about-dropouts-and-batchnormalization-in-cnn/

Brownlee, J. (2021, January 21). How to choose an activation function for deep learning. MachineLearningMastery.com. https://machinelearningmastery.com/choose-an-activation-function-for-deep-learning/

Vishwakarma, N. (2023, September 29). What is Adam Optimizer?. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2023/09/what-is-adam-optimizer/

Agrawal, S. K. (2023, September 29). Metrics to evaluate your classification model to take the right decisions. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/#:~:text=There%20are%20many%20ways%20for,used%20metrics%20for%20classification%20problems.

Nielsen, J. (2020, November 15). 10 usability heuristics for user interface design. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/