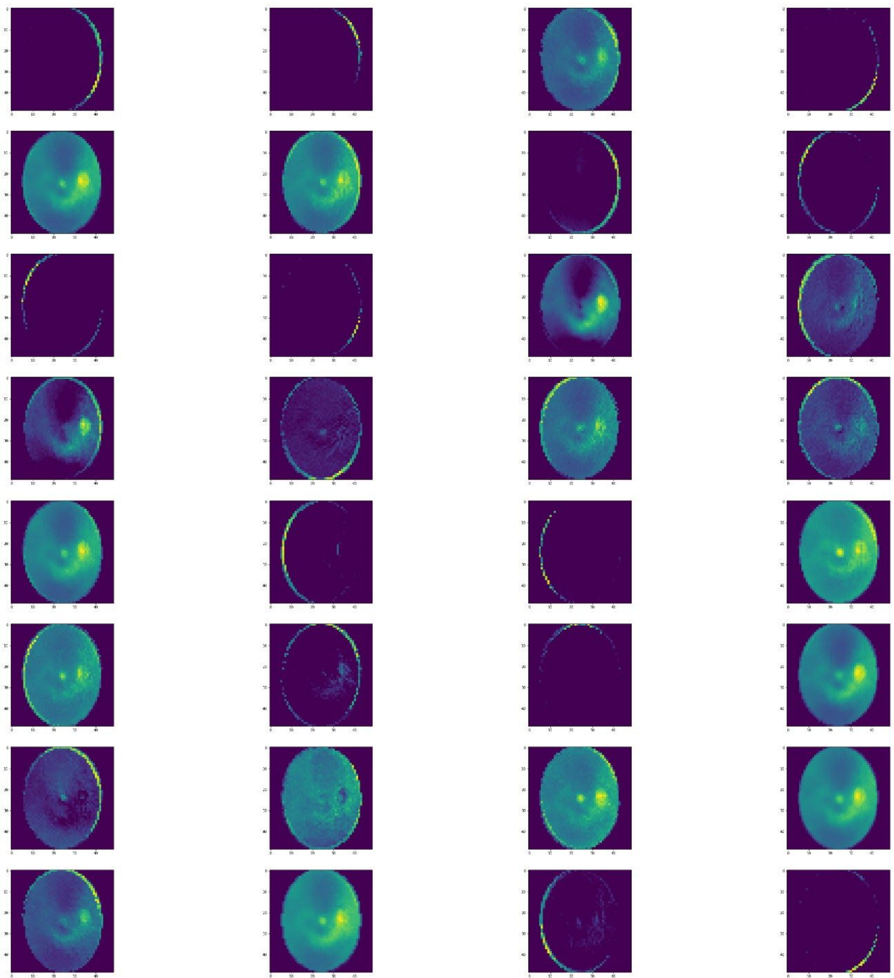
Keeping an Eye on CNNs

Data Analytics: Final Project

Group Incognito

Fall 2020



Esther Ko
Neli Davtyan
Ping Hsun Lee
Serra Topal Ismail Oglou
Eduardo Diaz Corona Reyes Retana

Introduction

"Every living person on this planet has their own unique pair of eyes. Each their own universe."
Ian Gray from movie "I Origins"

"I Origins" is an American movie telling a story of a molecular biologist who is researching the evolution of human eyes. He starts taking as many pictures of eyes as he possibly could for his study and uncovers evidence that fundamentally changes how society thinks. Though the movie is a fantasy/science fiction, it motivated us to think how much can be explored about a person through his eye pictures in today's reality. Living in the age of AI, Machine Learning, and having an open access to various databases provides an excellent opportunity to apply our skills in making people's life much easier and healthier.

Ocular diseases are a widespread problem around the world. Following the health statistics reports, even in developed countries such as the US, having diseases like cataract, glaucoma, age-related macular degeneration, diabetic retinopathy is quite common for people age 40 and older[1]. Why is it then that only half the people get annual eye exams? Ocular exams are not cheap, and are often not included in basic health coverage plans, therefore the cost associated is a key deterrent for many people. The shortage of ophthalmologists is yet another challenging factor for developing countries[2]*(Appendix 1).* The absence of regular checkups goes beyond affecting people on an individual level, it also dampens the government's ability to plan ahead for upcoming years and inhibits medical professionals' efforts in early detection research[3].

Due to the aging population, the number of blind and visually impaired people in the United States is predicted to double by 2030 and triple by 2050[4]. While understanding the importance of early diagnosis of visual defects (when treatment is most effective), and seeing the challenges of having widespread regular eye exams, in this paper, we will provide a solution to the problem by applying a machine learning model that will catch a person's vision defects from his fundus photographs.

Data

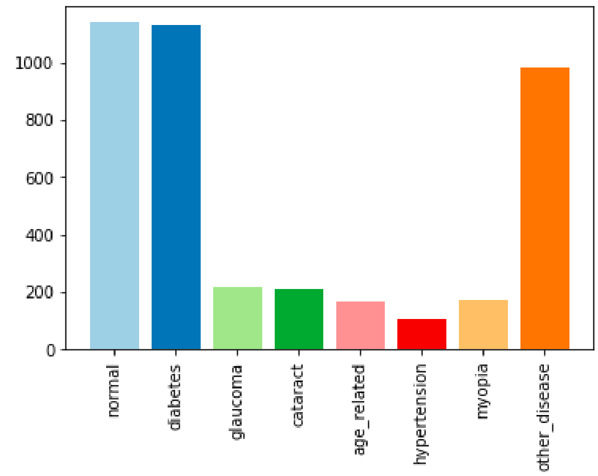
For our analysis we are using Ocular Disease Recognition (ODR) dataset from Kaggle database, which includes 3500 patient information collected by Shanggong Medical Technology Co., Ltd. from different hospitals/medical centers in China[5]. The dataset includes information on patient's age, sex, color fundus photographs from left and right eyes (captured by various cameras in the market, such as Canon, Zeiss etc.) and doctors' diagnostic keywords from doctors. Using the latest, the patients are classified into 8 groups, based on their specific ocular

[1] https://www.aao.org/newsroom/eye-health-statistics
[2] https://www.ncbi.nlm.nih.gov/pmc/articles
[3] https://www.cdc.gov/visionhealth/projects/economic_studies.htm
[4] https://www.cdc.gov/visionhealth/resourceshealth
[5] https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k?select=ODIR-5K

disease: Normal (N), Diabetes (D), Glaucoma (G), Glaucoma (G), Cataract (C), Age related Macular Degeneration (A), Hypertension (H), Pathological Myopia (M), Other diseases/abnormalities (O) *(Appendix 2)*.

Descriptive Statistics

The dataset contains information about 3500 patients, with mean age of 58, ranging from 1 to 91. The proportion of male and female patients is around the same - 54% and 46% respectively.



From the bar chart we can see that the leading eye disease is diabetes followed by other diseases. Glaucoma, cataract, age related Macular Degeneration, Hypertension and Myopia are almost equally common in the tested patients. Around one third of patients have normal vision, which based on the notations means that they have normal fundus both for right and the left eyes.

While looking at the differences between left and right eye defects, we have a balanced dataset. Approximately 11% of the patients have problems with only their left eye, 12% only with the right. For the 48% of the cases, people suffer from diseases of both eyes, and the rest of 29% is healthy.

Convolutional Neural Networks

A Convolutional Neural Network is a Deep Learning algorithm that takes in an input image, assigns importance to various parts in the image and is able to differentiate between them. CNNs take two sets of information, the input data and the kernel, to produce feature maps[6] *(Appendix 3)*.

Convolution is done as the kernel, which takes nearby pixels together, slides over the input. At every stop, the values on the kernel are multiplied by the value on the input and the results of this are summed up onto local outputs. Multiple kernels with identical dimensions are used to extract different attributes from an image and the outputs of this step are the feature maps[7] *(Appendix 4)*.

Feature mapping is followed by max pooling, which downsamples the feature maps to reduce its dimensionality. The maximum number will be chosen with the given kernel size onto a pooled

[6] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks
[7] https://medium.com/@aggirma/part-1-convolutional-neural-network-in-a-nutshell-89f81a329ec3

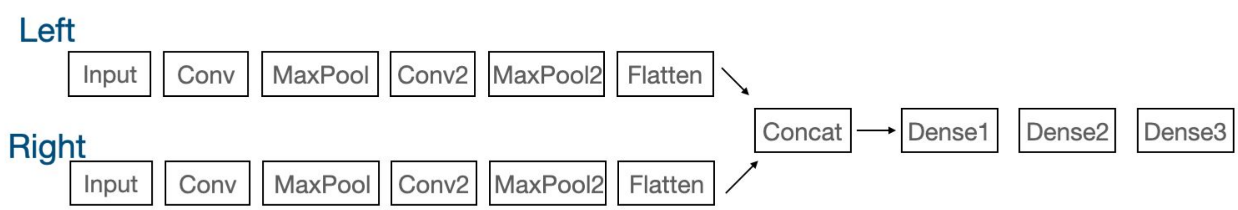
feature map. Before pooling is applied, feature maps are passed through non-linear activation functions and the resultant layers come together to form the feature learning. In the last step, feature maps form a fully connected layer after being flattened into a single column vector. This brings down the 3D map into the number of classes in the problem and the network can then predict the right class out of the given inputs.

CNNs are capable of extracting positional relationships in the image and construct several feature map abstraction layers. CNNs assumes that there exists a sense of locality, meaning that inputs that are close in proximity are related. This attribute enables it to capture the spatial relationship in the image whereas in a fully connected neural network such information is lost.
Additionally, CNNs have a faster training time, require less pre-processing, and have a lower computational cost as they need fewer parameters to learn in comparison to other classification algorithms[8].

CNN Model 1

-Using Keras Functional API for multiple input

1. Load in the image as 3-dim RGB
2. Run over 3 process before building fully connected neural network
   - Convolution2D: create feature map for image
   - MaxPooling2D: downsampling the feature map
   - Flatten: turn the layer to 1 column vector to be used by fully connected layer



-Accuracy for the model: 0.3771

CNN Model 2

For our second approach, we turned the image to grayscale to see if it can improve accuracy. Using the same model consisting of 2 convolution layers and merging the information of both eyes into one output of disease. This model didn't prove a significant change on the accuracy:
-Accuracy for the model: 0.3810

CNN Model 3

For our third approach, we changed the images into left and right eyes separately as the input, and used the labelled disease for each eye as the result. In this case, we can filter out the noises and unwanted information generated from combining two images as the input. The model improved the accuracy by almost 10%.
-Accuracy for the model: 0.4793

---

[8] https://medium.com/@aggirma/part-1-convolutional-neural-network-in-a-nutshell-89f81a329ec3

However, as the dataset has the following distribution of labels:

N	D	O	C	G	A	M	H	Total
2,873	1,608	708	293	284	266	232	128	6,392

That is, if the model simply outputs all the labels as ‘N’, we can get a 44.9% of accuracy, proving that our model still has a huge potential to grow.

CNN Model 4

Using the same logic of Model 2 we created a variation of Model 3 by changing the input images to black and white and therefore not having to deal with 3 dimensions as input. The logic behind this approach is that sometimes the colors create more noise in the model and therefore do not add any valuable information. However, in this case we found that the Grayscale Model actually decreased the overall accuracy by 3%.
-Accuracy for the model: 0.44

CNN Model 5

To improve our accuracy, we reduced the number of the labels. As the dataset contains 8 labels of diagnoses, we thought it might be too difficult for our model to learn. We simplified the classification to normal vs. abnormal, where abnormal representing the 7 other diagnoses. The dataset has the following distribution of labels: Normal (44.9%), Abnormal (55.1%). In this case, we used the same model as above for this approach and achieve:
-Accuracy for the model: 0.5776

Conclusion

Some possible explanations for this:

1. The images of the dataset are very similar so a more sophisticated model is needed in order to detect the differences
2. Some preprocessing of the images could enhance significant characteristics
3. For this particular problem, because we covered a wide range of disease, a larger data set or a model with many more layers could increase the accuracy

To further explore this, we use a VGG19 model that we found online to train on our binary classification data. In 10 epochs, we achieved a 63.29% accuracy and see no sign of overfitting. This proves that we can still improve our original model to better fit the data.

APPENDIX

Appendix 1

Table. Ophthalmologist Productivity by Country[a]

Country	Population in Millions[11]	Cataract Operations/y	CSR	No. of Ophthalmologists	Per Ophthalmologist: Population	Per Ophthalmologist: Cataract Operations[b]	Productivity, %[b]
Brazil	199.7	580 000	2900	4194	47 619	138	18
Bolivia	10.2	9500	930	133	76 981	72	10
Colombia	47.5	83 000	1752	1796	26 455	46	6
Cuba	11.6	31 000	2708	1244	9328	25	3
Ecuador	14.9	26 000	1737	453	32 906	57	8
México	116.4	200 000	1720	4190	27 778	48	6
Paraguay	6.6	9000	1350	103	64 390	88	12
Perú	29.7	47 000	1600	677	43 844	69	9
Venezuela	29.7	108 000	3650	1402	21 190	77	10

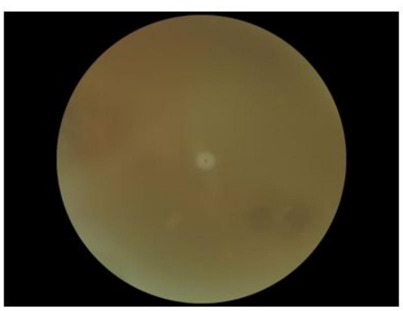
Abbreviation: CSR, cataract surgical rate.

[a] Performance and data on ophthalmologists courtesy of Van C. Lansingh, MD, PhD, Latin America, International Agency for the Prevention of Blindness, Weston, Florida (February 2012).

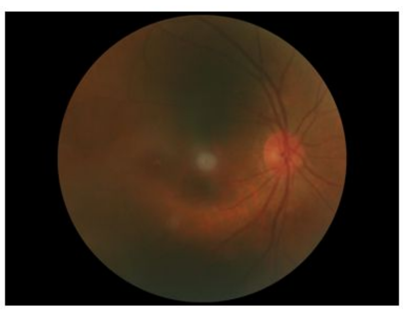
[b] Target productivity for cataract operations per ophthalmologist per year is 750 (100% productivity).

Appendix 2

	ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords	N	D	G	C	A	H	M	O
0	0	69	Female	0_left.jpg	0_right.jpg	cataract	normal fundus	0	0	0	1	0	0	0	0
1	1	57	Male	1_left.jpg	1_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0
2	2	42	Male	2_left.jpg	2_right.jpg	laser spot，moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	1
3	3	66	Male	3_left.jpg	3_right.jpg	normal fundus	branch retinal artery occlusion	0	0	0	0	0	0	0	1
4	4	53	Male	4_left.jpg	4_right.jpg	macular epiretinal membrane	mild nonproliferative retinopathy	0	1	0	0	0	0	0	1

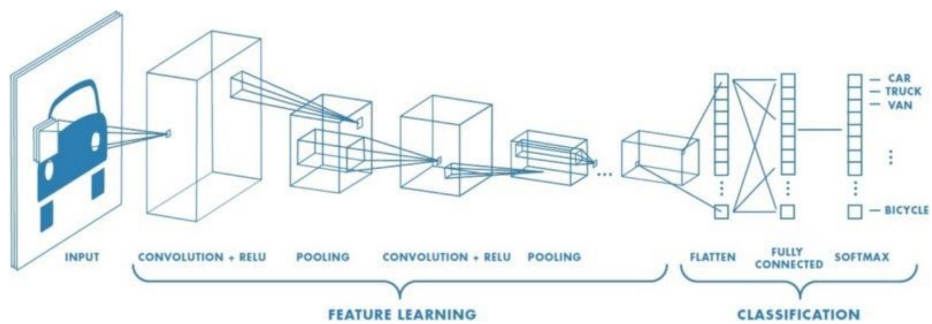


0_left.jpg: Cataract



0_right.jpg: Normal Fundus

Appendix 3



Appendix 4