

Capstone 1: Glaucoma Study and Detection

Problem: Glaucoma is a general term for a number of eye conditions that progressively damage the optic nerve, consequently causing vision loss. Diagnosis of glaucoma is complex, but is often associated with elevated intraocular pressure, optic nerve damage, and reduction in visual acuity and visual field¹. Vision loss from glaucoma is permanent, but progression may be slowed or halted through early diagnosis and treatment.

Client: Cambia Health is looking to hire an outside consultant use Cambia's collection of retinal images to help develop a machine learning algorithm to ultimately reduce medical practitioner time diagnosing patients with Glaucoma.

Data: For this project I will use data gathered from the Harvard Dataverse collected and uploaded by Ungsoo Kim². The retinal image dataset consists of 788 'normal control' images, 289 'early stage' images, and 467 'advanced stage' images. The images are already preprocessed, (scaled to 800 pixels and cropped so the nerve endings are of 240 pixels) thus ready to be used for machine learning tasks. However, this dataset will present itself particularly challenging as the number of retinal images needed for this task are less than desirable; thus representing a real-world problem of its own.

Approach: Time permitting I will use code to increase the size of the dataset with additional images by adding rotation and distortions to images, thereby hopefully increasing accuracy and generalizability of the model I intend to build. I will be building a Convolutional Neural Network model using Keras. Because the dataset is of smaller in size I will avoid the implication of building a model that can diagnose Glaucoma stages and simply create a model that can classify images as either having Glaucoma or not, leaving the challenge to the practitioner to decide stage and severity of disease.

Deliverables: I intend on building a small report using CDC data to describe Glaucoma to present the challenge, followed by a demonstration of the model classifying an image within Sagemaker. The code, Docker file, Sagemaker files, and Jupyter Notebooks will be available on Github. I intend to write a blog about the process as well as provide a short video presentation.

¹ American Optometric Association. Glossary of Common Eye & Vision Conditions website. <https://www.aoa.org/patients-and-public/eye-and-vision-problems/glossary-of-eye-and-vision-conditions>. Accessed June 7, 2018.

² Kim, Ungsoo, 2018, "Machine learn for glaucoma", <https://doi.org/10.7910/DVN/1YRRAC>, Harvard Dataverse, V1