Machine Learning Model For Glaucoma Detection: Potential Impact on Socio-Economically Disadvantaged Groups

Francis C. Chianu - 666560

Ishimwe Abedinego - 666221

School of Science and Technology, United States International University - Africa

SWE 2020: Machine Learning

Dr. Edward Ombui

July 31, 2024

TABLE OF CONTENTS

1. Introduction

The Sustainable Development Goals (SDGs), established by the United Nations, aim to address the main challenges that continue to persist all around the globe and promote positive change across various sectors. Among these goals, “Good Health and Well-being” stand out as a critical area of focus with far-reaching implication, particularly in regions subject to glaring health disparities. Within the context of Africa, achieving this goal is compounded by numerous challenges, including financial constraints, under-resourced healthcare systems, and limited access to promising medical solutions. Many individuals across the continent suffer from serious health conditions but face barriers to obtaining proper care due to socio-economic limitations and the inadequacies of existing healthcare infrastructures. The scarcity of resources and the high cost of medical services often result in delayed diagnoses and inadequate treatment, further exacerbating health disparities.

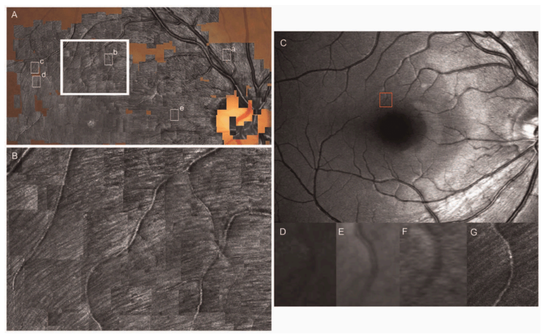
Using deep learning for disease detection with medical imaging provides a cost-effective way to screen for diseases. This method can mitigate the societal burden of diseases and reduce health disparities among different demographic groups. It can be applied in primary care settings and pharmacies, thus avoiding the need for patients to visit more expensive and crowded specialty clinics. In the field of ophthalmology, deep learning can simplify the detection of conditions such as glaucoma, a condition marked by increased pressure inside the eye that leads to vision loss. Glaucoma is characterized by elevated intraocular pressure causing changes in the optic disc and optic nerve, which poses a significant threat to vision health

1.1 Problem Statement

According to the World Health Organization (WHO), approximately 217 million people worldwide suffer from moderate to severe vision impairment. Glaucoma is a leading cause of irreversible blindness in individuals over 40 years old. Diseases like glaucoma, disproportionately affect racial and ethnic minorities and socioeconomically disadvantaged groups. Glaucoma, in particular, is more severe among Black and Hispanic populations.

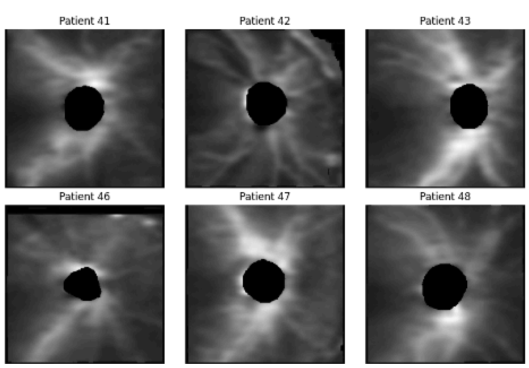
Currently, glaucoma is diagnosed through various methods, including a dilated eye exam to check for optic nerve damage, and an ocular pressure test to measure intraocular pressure, as elevated pressure may indicate glaucoma. These assessments assist ophthalmologists in determining the presence and severity of the disease. However, glaucoma often remains asymptomatic until it reaches an advanced stage, making regular checkups essential, driving up the cost and amount of labor.

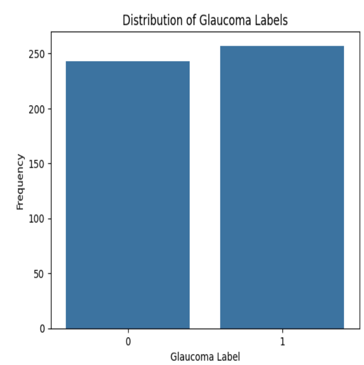
We propose a system that utilizes deep learning to analyze a dataset of labeled Retinal Nerve Fiber Layer Thickness (RNFLT) images to detect the presence of glaucoma. RNFLT images are obtained from Optical Coherence Tomography (OCT) scans and provide detailed cross-sectional views of the retina, specifically measuring the thickness of the nerve fiber layer. These measurements are crucial for diagnosing glaucoma, as thinning of this layer can be an early indicator of the disease. By leveraging RNFLT images, our predictive data analytics approach aims to enhance early detection and streamline the diagnostic process. The detailed structural information captured in these images will be captured through the layers of our model and will allow our deep learning model to identify subtle changes associated with glaucoma.



2 Design and Implementation

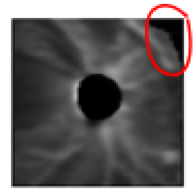
Unfortunately, finding a large dataset for RNFLT images proved to be very difficult. We used a dataset of 500 labeled RNFLT images from Harvard Medical School. This dataset was collected from 500 unique patients with and without glaucoma.





2.1 Background

Although advances in diagnostic methods like optical coherence tomography (OCT) and standard automated perimetry have greatly improved the detection and monitoring of ocular diseases, these techniques can produce data that is inherently prone to artifacts and noise, leading to inaccurate and unreliable data. Segmentation failures due to poor image quality or image defects undermine the clinical value of RNFLT. Deep learning models rely on clean, accurate and high-quality data for training and evaluation. The large number of noisy patient data challenges how effectively these models can learn from the data.



With that in mind, the primary focus was on cleaning, repairing, or reconstructing diagnostic images to remove artifacts and dark spots and to enhance the quality of our dataset. By addressing artifacts and noise, we ensure that the model trains on high-quality, accurate data, leading to improved performance, more reliable predictions, and a more effective diagnostic tool.

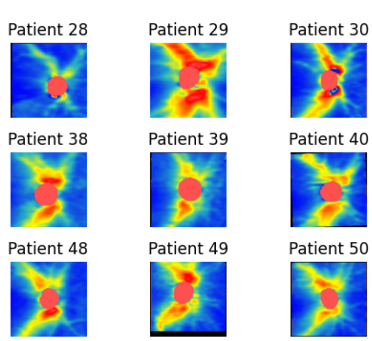
2.2 Data Preprocessing

The goal of the preprocessing tasks was to enhance the quality of the RNFLT images, and make them more suitable for training our model. This involved image cleaning and repair, image augmentation, 3D mapping, and batch processing.

- Image Cleaning and Repair:
    a. Artifact Removal - Images were cleaned to remove artifacts and dark spots, which involved clipping values, applying morphological operations, and using median and Gaussian blurring.
    b. Inpainting - Dark spots and missing areas were repaired using inpainting techniques, in order to improve image completeness and reduce potential biases caused by missing data.

- 3D Mapping: A colormap was applied to convert 2D images into 3D representations. This step helped in visualizing different intensity levels more effectively, to ensure better feature extraction from the mode during training.
- Image Augmentation: Various augmentations, including random cropping, rotation, and zooming, were applied to increase dataset diversity. This was aimed at improving the model's generalization by exposing it to a wider range of image variations.
- Batch Processing:
    a. Batch Handling - The dataset was processed in batches, which included cropping margins, resizing images, and ensuring consistent shapes. This approach was used to combat resource intensive preprocessing tasks.
    b. Combining and Shuffling - Processed batches were combined, and the dataset was shuffled to ensure randomness and reduce potential biases. Additionally, a testing dataset was extracted from the original dataset for model evaluation. This ensured that the training data was ready for training.

After image preprocessing, we ended up with much clear RNFLT image. There were still images with dark spots on the edges. However, this was a result of our team not properly handling images after they were augmented.



2.3 Model Overview

The model is a Convolutional Neural Network (CNN) designed for binary classification tasks. It processes 256x256 pixel RGB images to classify them into one of two categories, glaucoma or no glaucoma. Our model falls into the category of **predictive data analytics**. We aim to use historical image data from patients with and without glaucoma to identify patterns which we can use to predict the likelihood of glaucoma.

2.3.1 Leveraging Convolutional Neural Network

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed to process grid-like data, such as images. Their architecture and functionality make them particularly well-suited for medical image analysis, such as glaucoma detection in retinal images due to their proficiency in learning and extracting features from raw images. Additionally, CNNs employ hierarchical learning, with initial layers identifying basic features and deeper layers recognizing complex patterns. This combination enhances accuracy in detecting subtle variations and structural changes in retinal images.

2.3.2 Model Architecture

- Convolutional Layers: Three layers with 32, 64, and 128 filters, respectively, each followed by max pooling with a 2x2 pool size to reduce spatial dimensions and computational complexity while retaining essential features.
- Flatten Layer: Converts 3D feature maps into a 1D vector in order to transition from to fully connected layers.
- Fully Connected Layers: A dense layer with 128 units and ReLU activation is used for further processing of the features and help the model learn complex representations, followed by a dropout layer with a 0.5 rate to reduce overfitting.

- Output Layer: Final dense layer with 1 unit and sigmoid activation for binary classification.

2.3.3 Model Training

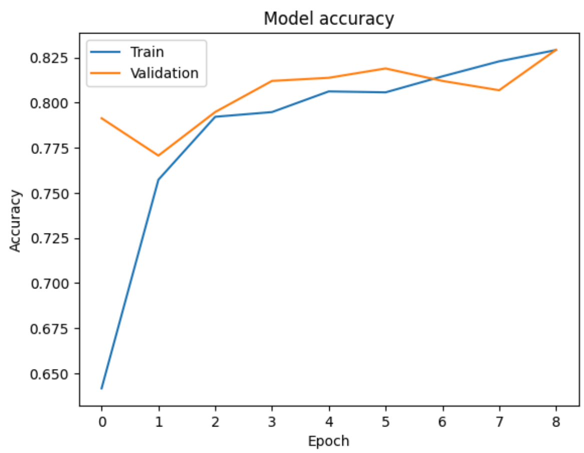
The model was compiled with the Adam optimizer (learning rate of 0.0002), binary cross-entropy loss for binary classification, and accuracy as the evaluation metric. The dataset was then divided into training and validation sets using a 4:1 ratio.

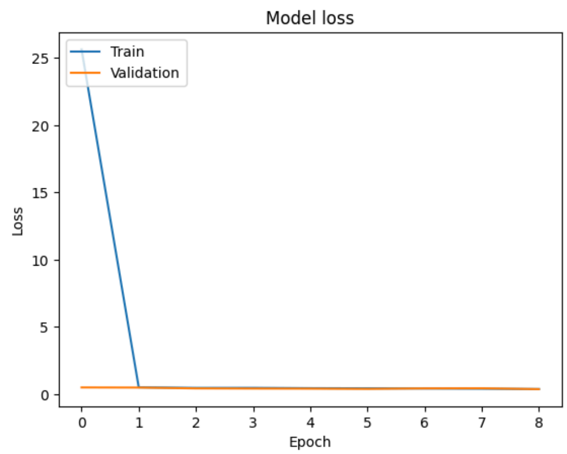
- The model was trained for 9 epochs.
- Training was done in batches of 32 samples.
- Validation data was used to monitor model performance on unseen data during training.

```
Epoch 1/9
63/63 [==============================] - 6s 71ms/step - loss: 14.4583 - accuracy: 0.6410 - val_loss: 0.5184 - val_accuracy: 0.7900
Epoch 2/9
63/63 [==============================] - 4s 64ms/step - loss: 0.4898 - accuracy: 0.7735 - val_loss: 0.4574 - val_accuracy: 0.7820
Epoch 3/9
63/63 [==============================] - 4s 63ms/step - loss: 0.4548 - accuracy: 0.7920 - val_loss: 0.4688 - val_accuracy: 0.7860
Epoch 4/9
63/63 [==============================] - 4s 66ms/step - loss: 0.4018 - accuracy: 0.8215 - val_loss: 0.4328 - val_accuracy: 0.7940
Epoch 5/9
63/63 [==============================] - 4s 65ms/step - loss: 0.3865 - accuracy: 0.8245 - val_loss: 0.4090 - val_accuracy: 0.8100
Epoch 6/9
63/63 [==============================] - 4s 64ms/step - loss: 0.3542 - accuracy: 0.8345 - val_loss: 0.4700 - val_accuracy: 0.7720
Epoch 7/9
63/63 [==============================] - 4s 66ms/step - loss: 0.3236 - accuracy: 0.8520 - val_loss: 0.4474 - val_accuracy: 0.8160
Epoch 8/9
63/63 [==============================] - 4s 67ms/step - loss: 0.3041 - accuracy: 0.8640 - val_loss: 0.4034 - val_accuracy: 0.8320
Epoch 9/9
63/63 [==============================] - 4s 65ms/step - loss: 0.2928 - accuracy: 0.8725 - val_loss: 0.4642 - val_accuracy: 0.8260
```

2.3.4 Fine Tuning

The previous versions of our model were overfitting, likely due to the combination of high data volume, numerous augmented images, and too many epochs. Initially trained for 25 epochs, we observed that after about 9 epochs, validation loss increased while training accuracy continued to rise, indicating overfitting. Furthermore to improve model performance, we modified the dataset by increasing the inpaint radius from 3 to 5, which resulted in sharper images and improved validation accuracy from 7.7 to 8.0. We also included a third convolution layer to the model, boosting validation accuracy to 8.2.





3 Model Evaluation

To evaluate my model, we used the test set we had extracted from the main dataset. The result we got showed an accuracy of 0.95 on the test dataset.

```
# Evaluate the restored model
loss, acc = new_model.evaluate(last_maps, last_labls, verbose=2)
print('Restored model, accuracy: {:5.2f}%'.format(100 * acc))
```

```
4/4 - 4s - loss: 0.1648 - accuracy: 0.9500 - 4s/epoch - 989ms/step
Restored model, accuracy: 95.00%
19/19 [==============================] - 25s 1s/step
(580, 1)
```

Although an accuracy of 0.95 indicates that our model performed very well on the test samples, it is difficult to conclusively claim that our model is as accurate as what is shown in the results. This is because we weren't able to procure a large enough test set from our original set with jeopardizing model training.

4 Challenges

During the development of the model, several key challenges were encountered:

- Limited Data Availability: The primary challenge was a small dataset of labeled retinal images, impacting the model's ability to generalize. We addressed this by using extensive data augmentation to artificially expand the dataset.
- Model Overfitting: Overfitting was a concern due to the small dataset, as the model tended to memorize rather than generalize. We employed regularization techniques such as dropout, early stopping, and hyperparameter tuning to combat this issue.
- Computational Resources: Training deep learning models is resource-intensive. Limited access to high-performance computing slowed the process and limited experimentation. We optimized the use of available resources to mitigate this challenge.

5 Potential Impact

Further steps in expanding and refining this model could significantly impact healthcare, notably in Africa, particularly concerning “Good Health and Well-being”. With improved model accuracy and a larger dataset to work with, the model could enhance the early detection of glaucoma, a critical need in areas with limited access to healthcare professionals. This advancement could enhance accurate early diagnosis which can lead to cost savings by preventing disease progression, contribute to better health outcomes and address disparities in healthcare access and quality.

References

AI for Medical Data Cleaning. Harvard Ophthalmology AI Lab. (2024, March 19). https://ophai.hms.harvard.edu/ai-for-medical-data-cleaning/

Harvard. (2024b, March 19). Home. Harvard Ophthalmology AI Lab. https://ophai.hms.harvard.edu/

Harvard-Ophthalmology-AI-Lab. (n.d.). Harvard-Ophthalmology-ai-Lab/EyeLearn: [Journal of Biomedical and Health Informatics 2023] artifact-tolerant clustering-guided contrastive embedding learning for ophthalmic images in glaucoma. GitHub. https://github.com/Harvard-Ophthalmology-AI-Lab/EyeLearn/tree/main

Hood, D. C., Bruna, S. L., Tsamis, E., Thakoor, K. A., Rai, A., Leshno, A., de Moraes, C. G. V., Cioffi, G. A., & Liebmann, J. M. (2022, February 23). Detecting glaucoma with only Oct: Implications for the clinic, research, screening, and AI development. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S135094622200012X?via%3Dihub

National Eye Institute. "Glaucoma." Accessed August 3, 2024. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/glaucoma#:~=Glaucoma%20is%20a%20group%20of,a%20comprehensive%20dilated%20eye%20exam.

Shi, M., Tian, Y., Elze, T., & Wang, M. (2024, February 29). RNFLT2VEC: Artifact-corrected representation learning for retinal nerve fiber layer thickness maps. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S1361841524000355?via%3Dihub