ConvNeXts and Vision Transformer Based Framework for Glaucoma Screening

Abdul Qayyum[1] and Moona Mazher[2] Imran Razzak[3]

[1] Université de Bourgogne, France
[2] University Rovira i Virgili, Spain
[3] University of New South Wales, Sydney, Australia

Abstract. Glaucoma is one the leading cause of blindness which can be prevented if diagnosed early. It is an irreversible condition and unfortunately, there is no cure. Glaucoma tends to happen slowly, often with no noticeable changes until after the damage is done. Recently, AIROGS challenge was organized with the aim to develop a robust framework for efficient glaucoma diagnosis. In this work, we present ConvNeXts and vision transformer based ensemble network for early diagnosis of glaucoma. Experiments are conducted on AIROGS dataset and results are available on Leaderboard (Team: mirazzak@deakin.edu.au).

Keywords: AIROGS · Glaucoma · Vision Loss

1 Introduction

Glaucoma is the leading eye diseases where vision is lost due to damage to the optic nerve and the connection between the eye and the brain. In Australia, 2 in 100 develop glaucoma in their life and 50% of them even do not know they are suffering glaucoma. If left untreated, it can lead to vision loss and blindnes. Initially, it starts with gradual loss of vision which reaches to considerable loss before we aware of this problem. It is an irreversible loss and has no cure. Early detection and treatment of glaucoma can help to delay or halt its progression. Thus, early diagnosis of Glaucoma has significant importance to avoid the vision impairment. Current practice for glaucoma includes physical examination of the eye that includes vision testing, peripheral visual field and microscopic examination of the retina, optic nerves head, arteries and veins. Furthermore special scans and photos of these structures are also performed to further investigate the type of glaucoma and appropriate treatment plan. However, manual screening is tedious, hence difficult for larger population especially developing and underdeveloping countries. Recently, AIROGS challenge was the key attempt with the aim to develop a robust framework for efficient Glaucoma diagnosis.Deep learning have been actively applied for medical image analysis and showed promising performance. In this work, we develop a ConvNeXts based framework for Glaucoma diagnosis. The **key contributions** of this work are

- present ensemble network using ConvNeXts and vision transformer for efficient Glaucoma diagnostic framework for early detection of glaucoma.
- the framework is simple and efficient in learning of depth-wise multiresolution features.

2 Methodology

In this work, we have used different pre-trained deep learning architectures such as DensNet201, Efficienet-B4, ConvNeXt and 2D vision transformer and fine-tuned on AIROGS dataset. We then ensembled different network using mean average to predict the score. We have considered different input size i.e. 224x224 and 384x384 for ConvNeXt and 2D vision transformer (VIT). The fully convolutional nature of ConvNeXts makes it simplest network and efficient in learning of depth-wise multiresolution features. As the dataset is highly imbalanced, Thus the class with lower number of samples requires more weights in comparison to the class with high number of samples. We have used cross-entropy loss function with loss weight distribution for higher class samples. For ungradibilty task, we have used the regret function to measure the gradibility or ungradibility. If the value of regret function is high, the samples are out of distribution and we marked as ungradbable. The formula to assign the regret function is given in Equation 1.

$$R = log\Big(\sum_{i}^{C} \frac{P_i}{P_i + P_i^{x_g^T}(1 - P_i)}\Big) \quad (1)$$

where P_i is the normalization factor and is the sum of the probabilities assignment of models that were trained with a specific value of the test samples. x_g^T is the projection that was computed using features from the trained model. C is the number of classes. The output ($O1$) is computed using binary classification prediction, and output ($O2$) was the binary score of gradeability prediction, output ($O3$) is the ungradable score that was calculated based on regret function. If score of regret function is high its ungradable else gradable.

3 Experiment

In work we present an efficient framework for Glaucoma diagnosis. We have performed experiment on AIROGS trainset which consist of 101,442 color fundus images from 54,274 subjects. We have used different pre-trained networks for glaucoma diagnosis. We have used Adam optimizer for training and set the learning rate 0.0001. Results on ensemble of different network showed that ensembler of ConvNeXt and vision transformer showed better performance than other networks.

4 Conclusion

In this work, we presented ensemble of ConvNeXt and vision transformer for Glaucoma detection. Experiments were conduced on AIROGS datasets. We have used the regret function to measure the gradibility or ungradibility. To deal with class imbalance problem, we have used cross-entropy loss function with loss weight distribution. Experiments are conducted AIROGS dataset and results on available on Leaderboard (Team: mirazzak@deakin.edu.au).

References

[1] de Vente, Coen, Vermeer, Koenraad A., Jaccard, Nicolas, van Ginneken, Bram, Lemij, Hans G., Sánchez, Clara I. (2021). Rotterdam EyePACS AIROGS train set (1.1.0) [Data set]. IEEE International Symposium on Biomedical Imaging 2022 (ISBI 2022), Kolkata, Calcutta, India.