Supplementary Information: A Two-Phase Deep Learning Approach to Glaucoma Estimation Using Retinal Layers Thickness

Linchuan Xu[1,*,+], Ryo Asaoka[2,3,4,*], Jiannong Cao[5], Hiroshi Murata[2,6], Yuri Fujino[2,3,4,7], and Kenji Yamanishi[1]

[1]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8656, Japan
[2]Department of Ophthalmology, The University of Tokyo, Tokyo, Japan
[3]Department of Ophthalmology, Seirei Hamamatsu General Hospital, Shizuoka, Japan
[4]Seirei Christopher University, Shizuoka, Japan
[5]Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[6]Department of Ophthalmology, National Center for Global Health and Medicine, Tokyo, Japan
[7]Department of Ophthalmology, Shimane University, Shimane, Japan
[*]unilin54092@gmail.com, rasaoka-tky@@umin.ac.jp,
[+]This work was done when Linchuan Xu was with The Hong Kong Polytechnic University.

1 Data Selection

Upon analysis, data with low reliability and duplicates were removed beforehand. In this study, the filtering process performed on the following bases:

- **Reliability**: The VF data with $FL \geq 0.33$ or $FP \geq 0.33$ or $FN \geq 0.33$ were removed. Also, the RT data with $SSI < 0.7$ were removed (where FL: Fixation failure, FP: False positive, FN: False negative, SSI: Signal strength index, are the criteria of ophthalmological measure test).
- **Duplicated data**: When a test for VF or RT was performed more than once a day, only the last one was selected.
- **Pairing RT and VF**: We study RT-VF pairs measuring the same glaucomatous loss. Therefore, it is expected that an RT and the corresponding VF be measured on the same day. However, owing to irregular diagnosis schedules, it is not common that a patient took both VF tests and OCT tests on the same day. To enlarge the data, we pair an RT and VF when the time difference between them is no more than six months.

The dataset that we used in the experiments cannot be published online because of ethical concerns.

2 Data Preprocessing

We resized RT from 512×128 into 224×224 using the resize method of OpenCV (https://pypi.org/project/opencv-python/) with interpolation=cv2.INTER_AREA. Afterwards, we re-scaled RT into the range $[0,1]$ and then standardized RT using mean $= [0.485, 0.456, 0.406]$ and standard deviation $= [0.229, 0.224, 0.225]$.

3 Training Setting

The algorithms were implemented with Pytorch V1.10.1 and were executed on one NVIDIA GeForce RTX 3090 GPU. We adopt ResNet-34 as our CNN backbone. We used pretrained ResNet-34 from Pytorch library. The training time for both the classification and regression took about 20 minutes.

We used SGD as the optimizer for the classification model. Weight decay was set to be $1e^{-4}$, momentum and nesterove are set to be true. Learning rate was initialized as $1e^{-2}$, and was divided by 10 after 80 epochs and 120 epochs using the multi-step learning rate scheduler. The total number of training epochs was 160. Batch size was set to be 16.

We used Adam as the optimizer for the regression model. Weight decay was set to be $1e^{-6}$, and learning rate was set to be $2.5e^{-4}$. The total number of training epochs was 160. Batch size was set to be 16.

We searched the value of η in $\{0.1, 0.2, 0, 5, 1, 2, 5, 10, 50\}$ via 5-fold cross-validation.