Wayfinding artificial intelligence to detect uncertain spots of OCT image in retinal diseases

Taiji Sakamoto (tsakamot@m3.kufm.kagoshima-u.ac.jp)
Kagoshima University Graduate School of Medical and Dental Sciences
Hideki Shiihara
Kagoshima University Graduate School of Medical and Dental Sciences
Shozo Sonoda
Kagoshima University Graduate School of Medical and Dental Sciences
Hiroto Terasaki
Kagoshima University Graduate School of Medical and Dental Sciences
Kazuki Fujiwara
Kagoshima University Graduate School of Medical and Dental Sciences
Ryoh Funatsu
Kagoshima University Graduate School of Medical and Dental Sciences
Ryosuke Shiba
Nidek (Japan)
Yoshiki Kumagai
Nidek (Japan)
Naoto Honda
Nidek (Japan)

Article

Keywords:

Posted Date: April 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1537787/v1

Abstract

The purpose of this study is to develop an artificial intelligence (AI) algorithm to identify spots with uncertain retinal structure using optical coherence tomography (OCT). SD-OCT B-scan images were classified into 189 normal and 111 disease-affected images. They were automatically segmented using a boundary layer detection model using deep learning. During segmentation, the AI model calculated the probability of the boundary surface of the layer for each A-scan. If this probability distribution is not biased toward one point, the layer detection is uncertain. This uncertainty was calculated by entropy and a value, referred to as the uncertainty index (UI), was calculated for each OCT image. The ability of the UI in classifying normal and disease images and the presence or absence of abnormalities in each layer of the retina were evaluated by area under curve (AUC). A heatmap that changes color according to the UI value was created. The UI of normal and disease-affected images (mean ± SD) were 1.76 ± 0.10 and 2.06 ± 0.22, respectively, showing a significant difference ($p < 0.05$). The AUC for the classification of normal and disease-affected images using UI was 0.93. The AUCs for the presence of abnormalities in each layer of the retina were: ILM: 0.588, NFL/GCL: 0.902, IPL/INL: 0.920, OPL/ONL: 0.882, EZ: 0.926, RPE/BM: 0.866. Our algorithm was able to identify retinal structural abnormalities on OCT with high accuracy, which may provide a basis for the diagnostic process as a new wayfinding AI method.

Introduction

The advent of deep learning has spurred remarkable developments in artificial intelligence (AI).[1] AI is also applied in the classification of various images in the field of ophthalmology with reports emphasizing the possibility of diagnosis based on images with high accuracy. Some studies have even started to work on clinical AI applications.[2–11] However, one major drawback of AI is its inability to show specific clinical findings on the basis of AI diagnosis from images; this is referred to as black box AI.[12] To date, heatmaps have been added to help understand to some extent the region of the fundus image referenced.[13, 14] However, this only complements AI diagnosis.

Recently, Adler-Milstein et al. proposed the idea of Wayfinding AI for next-generation diagnostics.[15] AI presents a final diagnosis right from the start. However, for medical professionals and patients, the journey of diagnosis starts by removing uncertain information from large amounts of data. Usually, it is through this journey that a care plan is developed. A final diagnosis right from the outset can never garner the trust of clinicians, emphasizing the need for wayfinding AI—a next-generation AI—that advises on the process, rather than the current AI that hands out final diagnoses.

Recently, a study found that, in the diagnosis of choroidal optical coherence tomography (OCT), cases that confused humans were equally confusing to AI.[16] However, the great benefit of AI is that it can quantify the level of confidence in the diagnosis and findings. Using this function, it is possible to numerically indicate the strength of each clinical finding as well as the final diagnosis. As the clinician can select the appropriate image from among those available, it is possible to induce the wayfinding process.

Thus, in this work, an algorithm is devised to identify abnormal sites by AI, based on the findings of OCT. Highly accurate segmentation retinal OCT images is possible using AI.[17–21] However, in reality, even humans have difficulty in segmenting because the layer structure in disease-affected eyes has already broken down. Therefore, the output of the model is ambiguous even in the artificial intelligence model. Paradoxically, by using the value of "ambiguity," the abnormal site can be quantitatively determined. In this study, the variation of the probability distribution calculated by deep learning from the automatic segmentation model of OCT B-scan images is quantified and applied to OCT anomaly detection. This, in itself, does not indicate a diagnosis, but rather is an indication of how far the image may be from normal. This characteristic sets conventional diagnostic apart from AI-based diagnostics. If abnormal parts can be identified using this algorithm, the trust of the clinicians would be secured, which is exactly what a wayfinding AI sets out to do, which is, the focus of this paper.

Results

After consultation with two experts (HS and SS. If diagnosis was split, the third expert HT was included) the images were classified into 189 normal and 111 disease-affected eyes. Thirty cases had abnormalities at the boundary between internal limiting membrane (ILM); 44 cases had abnormalities at the boundary between nerve fiber layer (NFL) and ganglion cell layer (GCL); 50 cases had abnormalities at the boundary between inner plexiform layer (IPL) and inner nuclear layer (INL); 48 cases had abnormalities at the boundary between outer plexiform layer (OPL) and outer nuclear layer (ONL); 78 cases had abnormalities in the ellipsoid zone (EZ); and 48 cases had abnormalities at the boundary between RPE and the Bruch's membrane (BM).

Relationship between the presence or absence of disease and overall uncertainty

The overall uncertainty index (UI) of the normal image was 1.76 ± 0.10, and the UI of the disease image was 2.06 ± 0.22, showing a significant difference ($p < 0.05$, t-test) (Table 1). **Figure 1** shows the receiver operating characteristic (ROC) curve for the classification of normal and disease images using UI. For this ROC curve, AUC = 0.92. It is suggested that it should be possible to classify diseased and normal eyes using UI.

Table 1

	UI of normal eyes (n=189)	UI of diseased eye (n=111)	p-value
All	1.76 ± 0.10	2.06 ± 0.22	< 0.001
ILM	1.67 ± 0.12	1.70 ± 0.14	0.15
NFL/GCL	1.91 ± 0.20	2.32 ± 0.29	< 0.001
IPL/INL	2.12 ± 0.23	2.64 ± 0.31	< 0.001
OPL/ONL	1.96 ± 0.29	2.42 ± 0.27	< 0.001
EZ	1.55 ± 0.22	2.12 ± 0.43	< 0.001
RPE/BM	1.58 ± 0.18	1.93 ± 0.32	< 0.001

UI: uncertainty index

Relationship between the boundary abnormality of each layer and UI

Figure 2 and **3** show the distribution map of the image classification of each layer of the retina and the ROC curve. The UI ROC curve for the presence or absence of ILM boundary abnormalities had AUC = 0.588. The UI ROC curve for the presence or absence of NFL/GCL boundary abnormalities had AUC = 0.902. The UI ROC curve for the presence or absence of IPL/INL boundary abnormalities had AUC = 0.920. The UI ROC curve for the presence or absence of OPL/ONL boundary abnormalities had AUC = 0.882. The UI ROC curve for the presence or absence of EZ had AUC = 0.926. The UI ROC curve for the presence or absence of RPE/BM boundary abnormalities had AUC = 0.866. The representative cases are covered in **Supplementary Figure S1**.

Heatmap creation

Case 1: A representative case of heatmap image using UI is displayed demonstrating the usefulness of the UI heatmap in the diagnosis of simple diabetic retinopathy (**Figure 4**). The subject was a 54-year-old male. Fundus images, OCT B-scan examination, and retinal thickness examination are consistent with simple diabetic retinopathy. However, on the heatmap (F) based on the uncertainty index of the image of the OPL/ONL boundary layer, a light blue island with a large uncertainty index is observed in the layer below the macula. With further magnification, the OPL/ONL layer is disturbed. The UI heatmap points out changes in the retina that one may overlook.

Case 2: The case involves a 78-year-old male. The patient is being treated for exudative age-related macular degeneration (**Figure 5**) based on the fundus photograph. In the UI heatmap image, there is a high UI area in the pigment epithelial layer, but this is the lesion site and is consistent with the diagnosis. However, in the UI heatmap of the ILM layer, there is a part where the UI is high. Then, looking at the OCT B-scan image of that part, it was found that the anterior retinal membrane was present, demonstrating that the UI heatmap points out lesions that are often overlooked in normal medical care.

Discussion

The existing automatic segmentation model of OCT B-scan images by deep learning enables the quantification of the certainty of the boundary surface determination of each layer of the retina. By applying the quantified data, it is possible to detect abnormal parts.

There have been reports of using deep learning to classify OCT normals and abnormalities with high accuracy.[22, 23] However, deep learning has a problem that the analysis process is a black box and it is difficult to clarify the reason for the classification.[12] This is a major obstacle to clinical application. In this study, abnormalities are detected by quantifying the analysis process, and it is possible to numerically determine why the abnormalities are classified as they are. In other words, the process can be rephrased as explainable AI.

The advantage of this method is that it is easy to understand the degree of abnormality because quantified numerical values can be obtained. In addition, as abnormalities can be detected for each layer, it is easy to examine which layer has an abnormality. We believe that the thought processes of the physician concerning these parts can be a companion toward the correct diagnosis rather than providing the final diagnosis. This fits the concept of wayfinding. Moreover, as the segmentation model is applied, it is not necessary to incorporate a new program for abnormality detection into the machine. Along similar lines, in the past, there has been a report of the OCT segmentation model being applied to an automatic diagnostic model.[24] However, in that report, transfer learning using a deep learning model was performed in the diagnostic process, and because the diagnostic process is a black box, it is very different from this model.

According to the AUC obtained in terms of specificity and sensitivity, the ability to detect abnormalities at the ILM interface was not high. The lesions on the ILM interface are mainly ERM, but in the OCT image of ERM, the brightness of only the lesion site increased. This is probably because the entropy value often remained low in the segmentation such that there are many cases wherein the boundary surface is not disturbed. In contrast, strong abnormalities were detected between the IPL/INL interface and EZ. Diabetic macular edema is a typical disease that causes IPL/INL abnormalities, and AMD and CSC are typical diseases that cause EZ abnormalities. However, it is probable that the entropy value was high and the detection power was high because of the significant changes in the OCT image due to the lesion.

In this study, the entropy of each layer of OCT B-scan in one cross section passing through the macula is averaged and calculated as UI. However, by applying this method to the volume scan of OCT and displaying the entropy of each measurement point on a color scale and mapping, the lesion site can be visually comprehended, as shown in Figs. 4 and 5 below. It is also possible to visualize the distortion in the retinal layer in the case of early diabetic retinopathy (Fig. 4). To identify unexpected lesions that are often overlooked in regular medical care, for example, when treating age-related macular degeneration, the doctor may concentrate on the macula while often overlooking mild ERM; however, as shown in Fig. 5, ERM can be diagnosed at an early stage. In this method, an area around the abnormal area is attached,

and the area is scanned to obtain an OCT B-scan image. With this software, the average entropy value of each layer is calculated, and it is possible to numerically indicate the layer with an abnormality. In other words, in the heatmap, the area where there is a high possibility of any abnormality is marked, and scanned with OCT, making it possible to indicate the cross section of the image that is abnormal.

The problem with this model is that it has difficulty identifying the disease. There are many AIs that provide a final diagnosis, but this is not the intent of this AI; and as such, this difficulty is not considered a major drawback. This AI pinpoints "abnormal" images in the OCT image that are otherwise invisible in the fundus image. Should a person specifically identify an overlooked or unnoticed abnormality, they would think about it and determine the next necessary test for the examination which would finally lead to a correct diagnosis. Wayfinding AI has yet to be clearly defined. However, the AI developed in this study is in line with the idea of Wayfinding AI in that reaching the correct diagnosis through trial and error helps human thinking, and this diagnosis process has been adopted for a long time. The proposed AI is a new AI that would be more acceptable to doctors than the normal AI, which remains a black box in the diagnostic process. Overdependence on AI may result in unthinkable misdiagnosis, which the proposed AI can prevent.

Methods

This study was approved by the Ethics Committee of Kagoshima University Hospital (Kagoshima, Japan) and registered with the University Hospital Medical Network (UMIN)-clinical trials registry (CTR). The registration title is “UMIN000031747, Research on retinal/choroidal structure analysis by novel image analysis technique and machine learning.” on March 2018. A detailed protocol is available at, https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000036250. A written informed consent was obtained from all of the subjects after an explanation of the procedures to be used and possible complications. All of the procedures conformed to the tenets of the Declaration of Helsinki.

The subjects underwent OCT at Sonoda Eye Clinic, Kagoshima, Japan from January 15, 2019, to March 29, 2019, with the data constituting OCT images of 617 eyes. OCT images of the macular region passing through the fovea centralis were analyzed. Images with poor quality because of the opacity of the optic media, poor fixation, etc. as well as images that were inverted due to posterior staphyloma were excluded. From the images that survived the exclusion criteria, 300 were randomly selected for the analysis.

Imaging Protocol

Imaging was performed using SD-OCT (RS-3000 Advance2) manufactured by NIDEK (Tokyo, Japan). OCT was taken horizontally through the fovea centralis. The image was extracted as a 1024 x 512 pixel Microsoft Windows bitmap image (bmp).

Segmentation AI

All OCT images were segmented using an OCT B-scan automatic image segmentation model that utilizes deep learning.[25] For a brief overview, the segmentation model is based on U-Net[26] and consists of an encoder and decoder, a skip-connection between the two, and a multiple dilated convolution (MDC) block. The input is an OCT image, and the output is a probability map for each boundary layer (**Figure 6**). The encoder and decoder perform 7×1 vertical convolution to extract features of the horizontal edges (vertical brightness changes). The MDC block expands the receptive field by combining convolutions with different dilations to capture the positional relationship of a wide range of features. In the output layer of the model, SoftMax is applied in the vertical direction (direction of the A-scan). This makes it possible to obtain a probability distribution of the position (depth) of each boundary layer from the A-scan. Finally, the position (depth) with maximum probability distribution in each A-scan is detected as the boundary layer. This model allows segmentation of ILM, NFL-GCL boundaries, IPL-INL boundaries, OPL-ONL boundaries, EZ, and RPE-BM boundaries.

Calculating uncertainty

Using the above automatic segmentation model, the certainty of the calculation of the retinal interface was calculated. In the model, of the 512 points, initially those corresponding to the boundary layer are calculated for each A-scan of the OCT image, whereby the probability distribution of the boundary layer is the output, and the point with the highest probability is actually detected as the boundary layer. If the probability distribution is not biased toward one point and varies, the detection of the boundary layer is considered to be uncertain. The variation in this probability distribution was calculated by entropy. Entropy is an index indicating the degree of chaos and irregularity of a state, and is calculated by the following formula.

The entropy value ($\mathrm{Entropy}_{x,l}$) at position x of the boundary layer l of the A-scan is

$$\mathrm{Entropy}_{x,l} = -\sum_{z} p_{x,z,l} \log p_{x,z,l}$$

Here, $p_{x,z,l}$ is the output value of the network at coordinates (x, z) of the boundary layer l.

In other words, the larger the entropy, the more the probability distribution is scattered and the more uncertain layer detection is. This entropy was calculated for each boundary layer of the retina in the OCT image with the average value defined as the UI. The lower the UI, the smaller is the variation in the probability distribution, and the higher the UI, the greater is the variation in the probability distribution (Fig. 7). An example of segmentation using UI is provided in Fig. 8.

Heatmap creation

In the proposed method, a heatmap for each layer can be created by arranging the entropy calculated for each A-scan in the OCT volume (**Figure 9).** The heatmap created was smoothed by a Gaussian filter with

σ = 1 and normalized so that the minimum value would be 0 and the maximum value would be 1. To colorize the normalized heatmap, jet colormap was applied.

Image labeling

Two retina experts labeled the presence or absence of abnormalities in each layer of the OCT image. Those with epiretinal membrane (ERM), retinal edema, hard exudate, retinal pigment ethitelium (RPE) abnormality, serous retinal detachment (SRD), pigment epithelial detachment (PED), and drusen were defined as abnormalities. The retina was also examined to determine in which layer these abnormal findings were located. If two examiners had different opinions, the abnormalities were established after discussions between the two examiners.

Statistical analysis

All statistical analyses were performed with SPSS statistics 19 for Windows (SPSS Inc., IBM, Somers, NY). The difference between the mean values of normal and abnormal UI was examined using the t-test. The ability to classify normal and abnormal parts was evaluated by the AUC of the ROC curve. A p-value of 0.05 or less was considered significant.

Data availability

All data generated or analysed during this study are included in this published article and its supplementary information files.

Declarations

Funding:

Grant-in-Aid for Scientific Research from the Ministry of Education,

Science, and Culture of the Japanese Government 21H03095 and 22K16975

References

1. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning.*nature* 521, 436-444 (2015).
2. Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama* 316, 2402-2410 (2016).
3. Abràmoff, M. D. et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigative ophthalmology & visual science*57, 5200-5206 (2016).
4. Gargeya, R. & Leng, T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 124, 962-969 (2017).

5. Ting, D. S. W. et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes.*Jama*318, 2211-2223 (2017).
6. Burlina, P. M. et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA ophthalmology* 135, 1170-1176 (2017).
7. Burlina, P. M. et al. Use of deep learning for detailed severity characterization and estimation of 5-year risk among patients with age-related macular degeneration. *JAMA ophthalmology* 136, 1359-1366 (2018).
8. Li, Z. et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*125, 1199-1206 (2018).
9. Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B. S., Waldstein, S. M. & Bogunović, H. Artificial intelligence in retina.*Progress in retinal and eye research* 67, 1-29 (2018).
10. Grassmann, F. et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* 125, 1410-1420 (2018).
11. Brown, J. M. et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks.*JAMA ophthalmology* 136, 803-810 (2018).
12. Castelvecchi, D. Can we open the black box of AI? *Nature News* 538, 20 (2016).
13. Keel, S., Wu, J., Lee, P. Y., Scheetz, J. & He, M. Visualizing deep learning models for the detection of referable diabetic retinopathy and glaucoma. *JAMA ophthalmology* 137, 288-292 (2019).
14. Wang, Z. et al. Artificial intelligence and deep learning in ophthalmology. A*rtificial Intelligence in Medicine*, 1-34 (2020).
15. Adler-Milstein, J., Chen, J. H. & Dhaliwal, G. Next-Generation Artificial Intelligence for Diagnosis: From Predicting Diagnostic Labels to “Wayfinding”. *Jama* 326, 2467-2468 (2021).
16. Sonoda, S. et al. Artificial intelligence for classifying uncertain images by humans in determining choroidal vascular running pattern and comparisons with automated classification between artificial intelligence. *Plos one* 16, e0251553 (2021).
17. Maloca, P. M. et al. Validation of automated artificial intelligence segmentation of optical coherence tomography images. *PloS one* 14, e0220063 (2019).
18. Kugelman, J. et al. Automatic choroidal segmentation in OCT images using supervised deep learning methods. *Scientific reports* 9, 1-13 (2019).
19. Pekala, M. et al. Deep learning based retinal OCT segmentation. *Computers in biology and medicine*114, 103445 (2019).
20. Breger, A. et al. Supervised learning and dimension reduction techniques for quantification of retinal fluid in optical coherence tomography images. *Eye* 31, 1212-1220 (2017).
21. Tian, J. et al. Real-time automatic segmentation of optical coherence tomography volume data of the macular region. *PloS one* 10, e0133908 (2015).

22. De Fauw, J. et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*24, 1342-1350 (2018).
23. Rim, T. H. et al. Computer-aided detection and abnormality score for the outer retinal layer in optical coherence tomography.*British Journal of Ophthalmology* (2021).25. 22
24. Xiancheng, W. et al. in Procedia Computer Science: *International Conference on Data Science* (ICDS 2018). 8-9.
25. Ronneberger, O., Fischer, P. & Brox, T. in International Conference on Medical image computing and computer-assisted intervention. 234-241 (Springer).
26. Lee, C. S., Baughman, D. M. & Lee, A. Y. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images.*Ophthalmology Retina*1, 322-327 (2017).

Figures

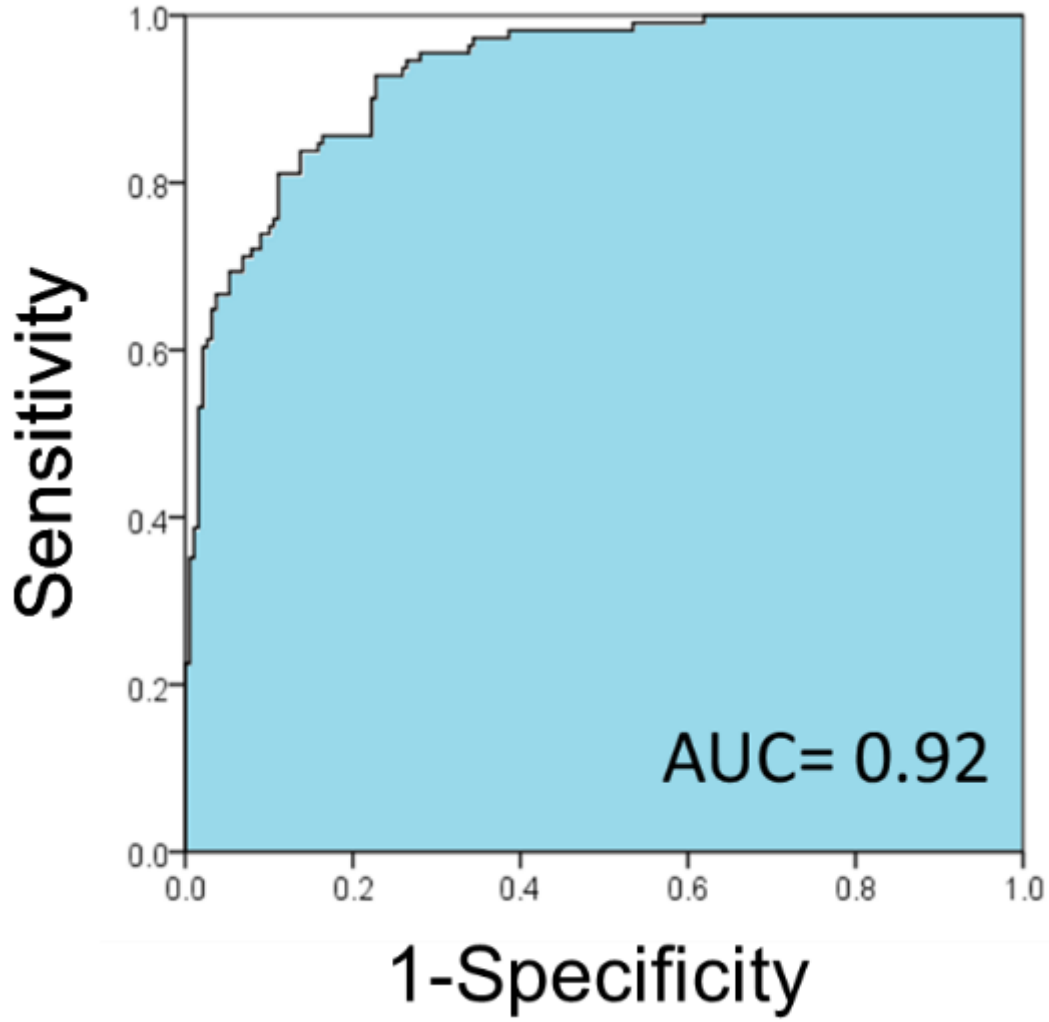

Figure 1

The ROC curve for the classification of normal and abnormal images by mean UI of each layer of the retina, classified by AUC = 0.92.

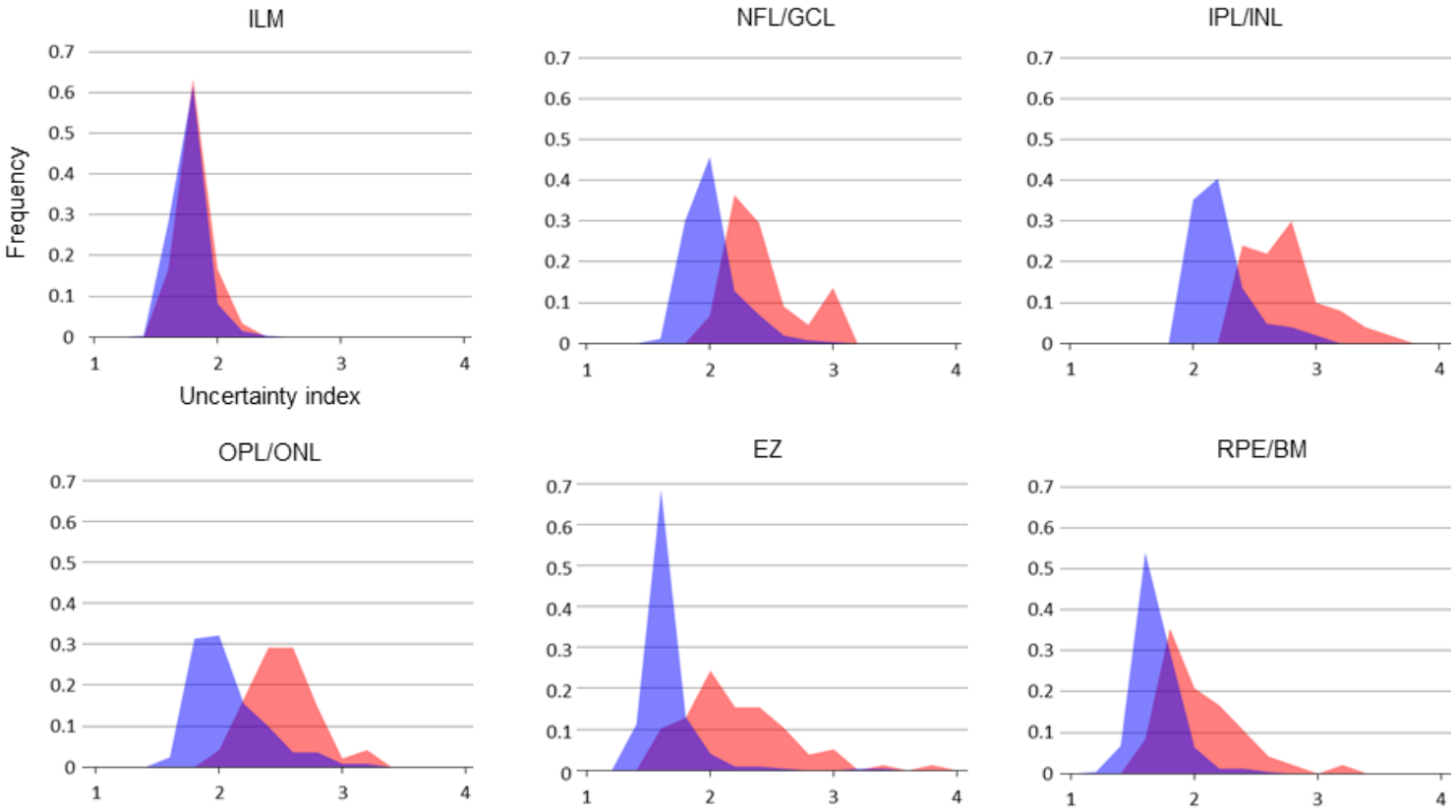

Figure 2

It was observed that IPL/INL and EZ can be classified with high accuracy.

Blue: Normal UI distribution map; Red: Abnormal UI distribution map

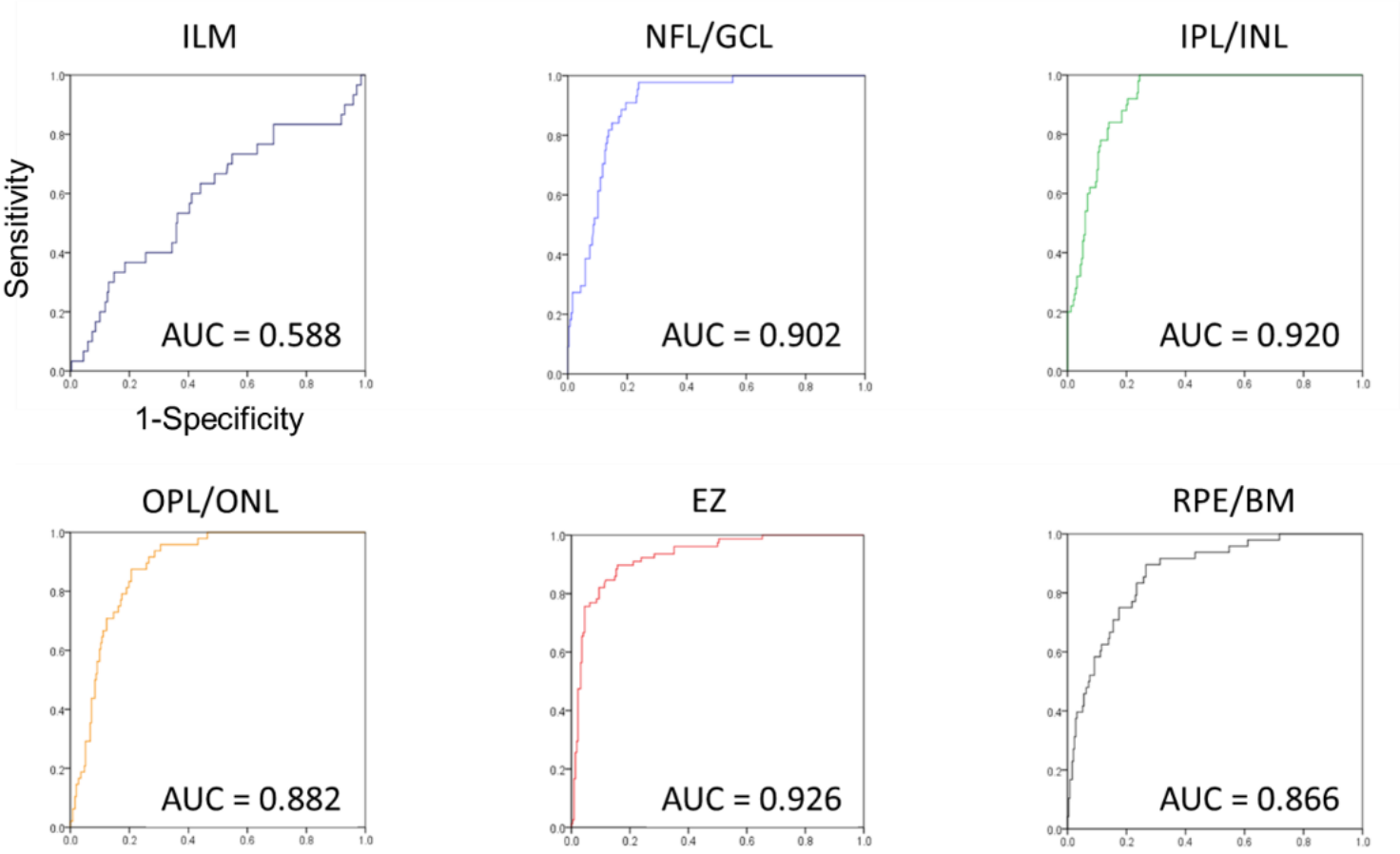

Figure 3

ROC curve for each layer of the retina with normal and abnormal classification by UI. Except for ILM, the normal and abnormal layers are classified with high accuracy.

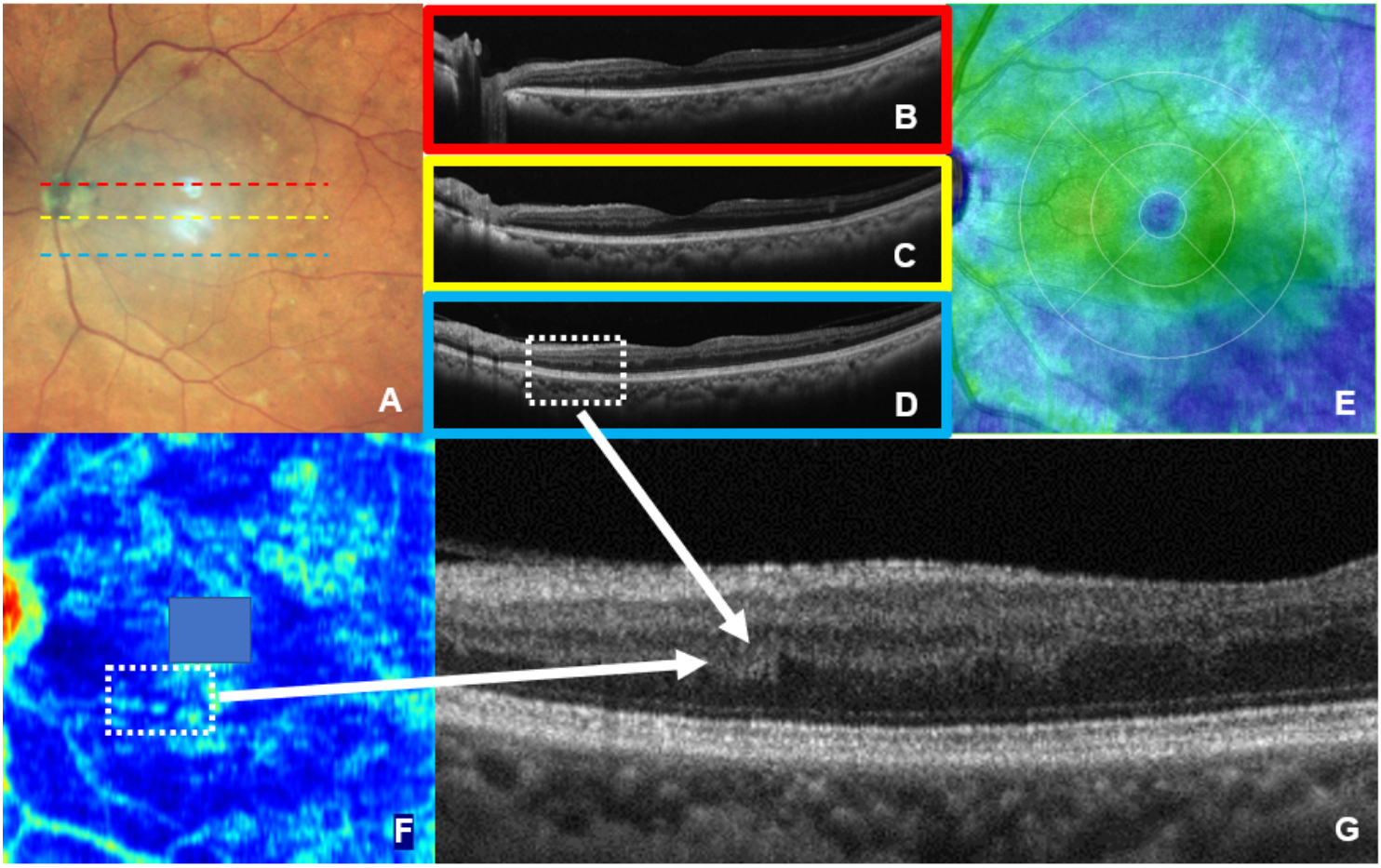

Figure 4

Case 1. Multimodal image of simple diabetic retinopathy. Color fundus photograph (A). OCT image of A-scan with dotted red, yellow, and blue lines (B, C, D). Color map (E) reflecting the thickness of the retina for the same case. These findings are consistent with the findings of simple diabetic retinopathy. However, when the heatmap (F) based on the uncertainty index shows the image of the OPL/ONL boundary layer, a light blue island with a large uncertainty index is observed in the layer below the macula. Therefore, if that part is magnified (dotted line qualification part of D and F), the presence of a distortion in OPL/ONL (G) can be observed . From this, it can be instantly determined that distortion has begun at the layer structure of the retina despite simple retinopathy. As the macula does not have a layered structure, the entropy value of the OPL/ONL boundary layer is often high even in normal eyes, and hence shielded to make it distinguishable.

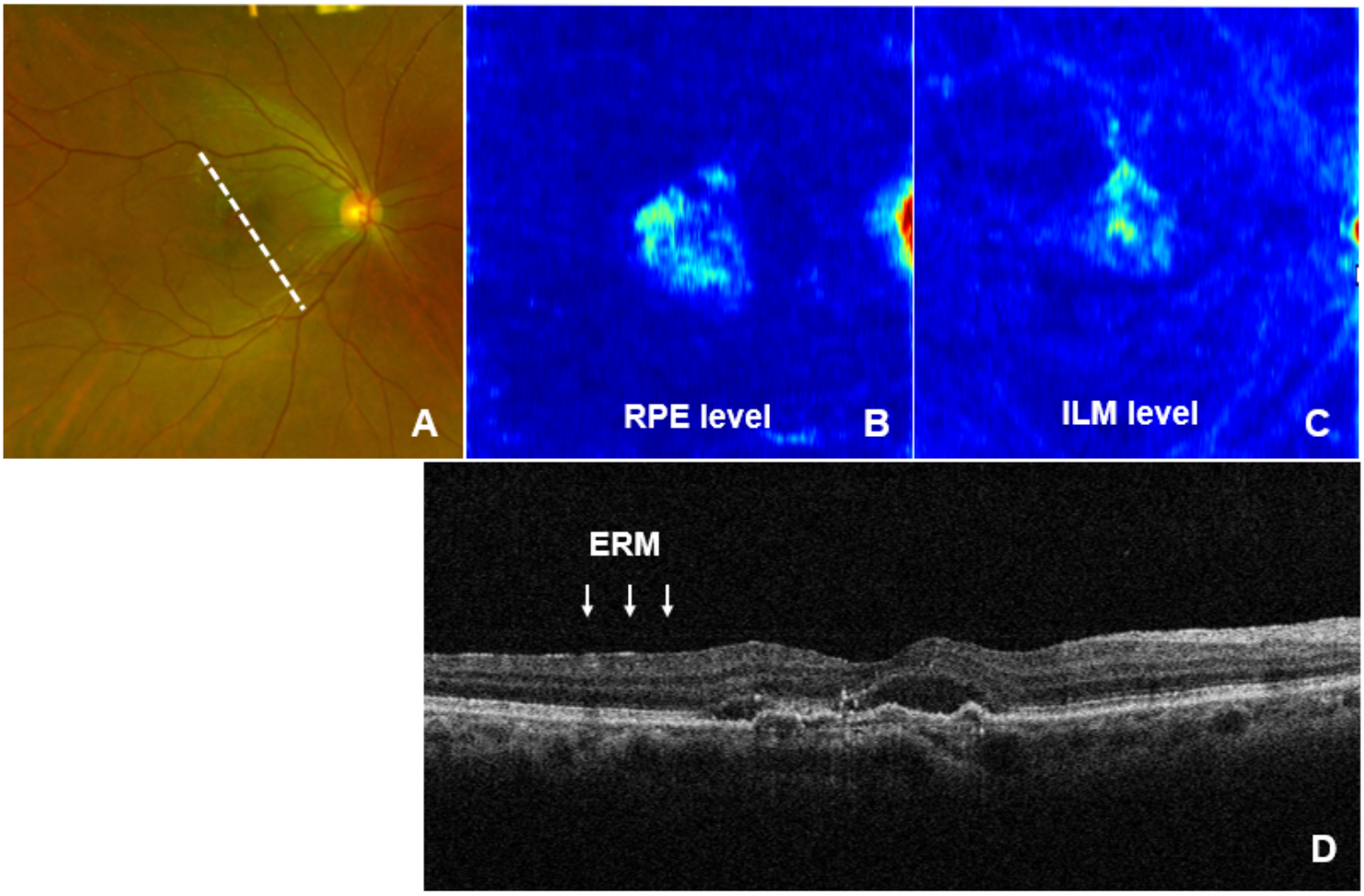

Figure 5

Multimodal image of exudative age-related macular degeneration. (A) Color fundus photograph. (B) RPE level UI heatmap image. (C) ILM level UI heatmap image. (D) OCT B-scan image, scan of the white part of the dashed line in (A). UI: uncertainty index, RPE: retinal pigment epithelium, ILM: internal limiting membrane, ERM: epiretinal membrane.

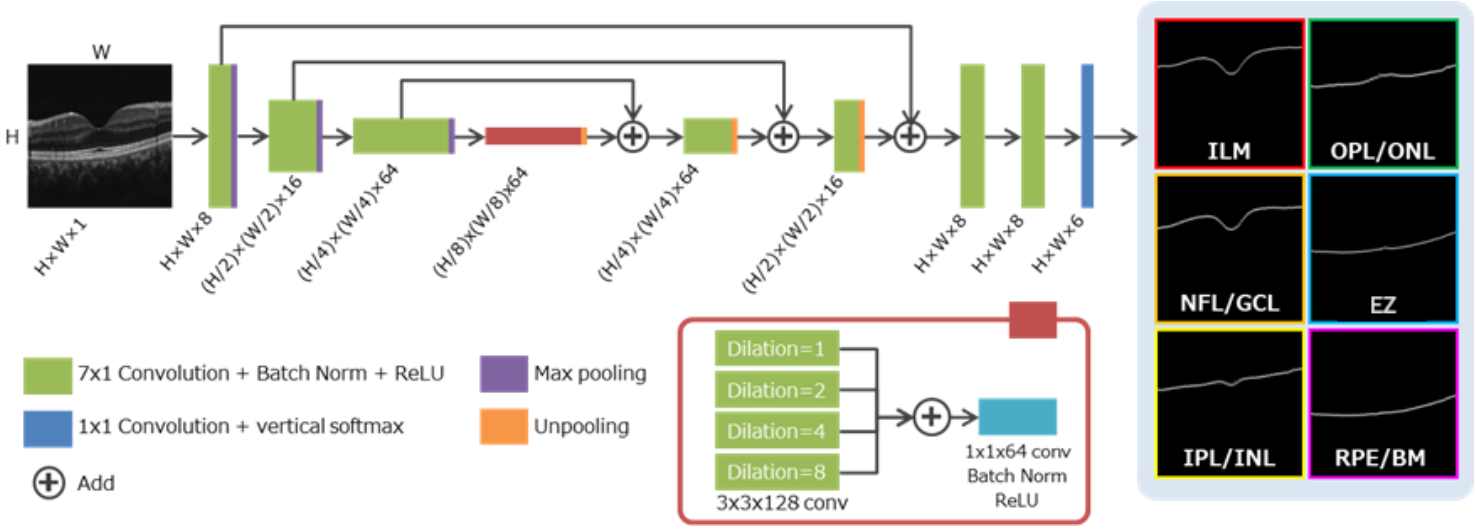

Figure 6

An OCT B-scan automatic image segmentation model was created based on U-Net. The output of this model is the probability distribution of each boundary layer, allowing the automatic detection of the boundary for each layer of the retina.

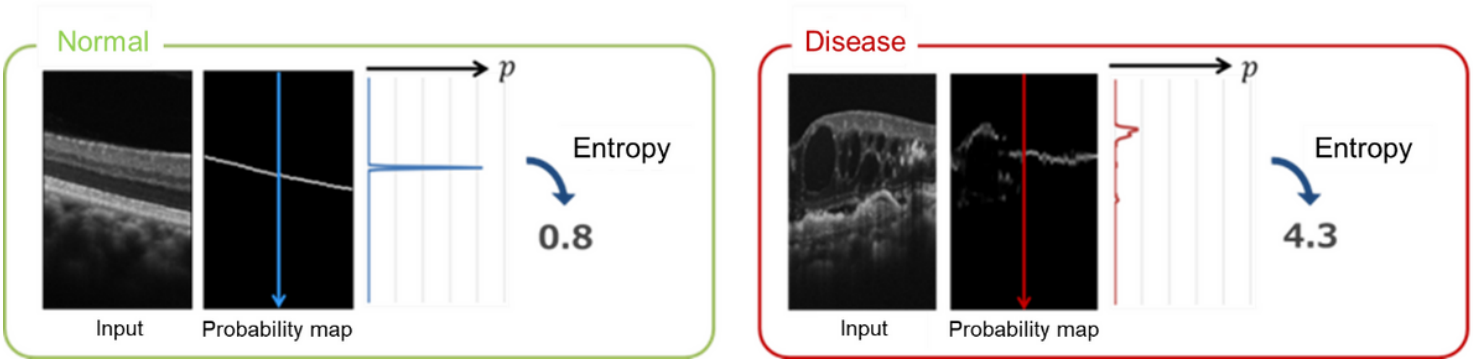

Figure 7

Conceptual diagram of the entropy value of each point in OCT A-scan. The OCT A-scan line in normal eyes is blue, and in disease-affected eyes, the scan line is red. In normal eyes, the probability distribution does not vary, so entropy is small, but in disease-affected eyes, the variation is large, so the entropy value is large.

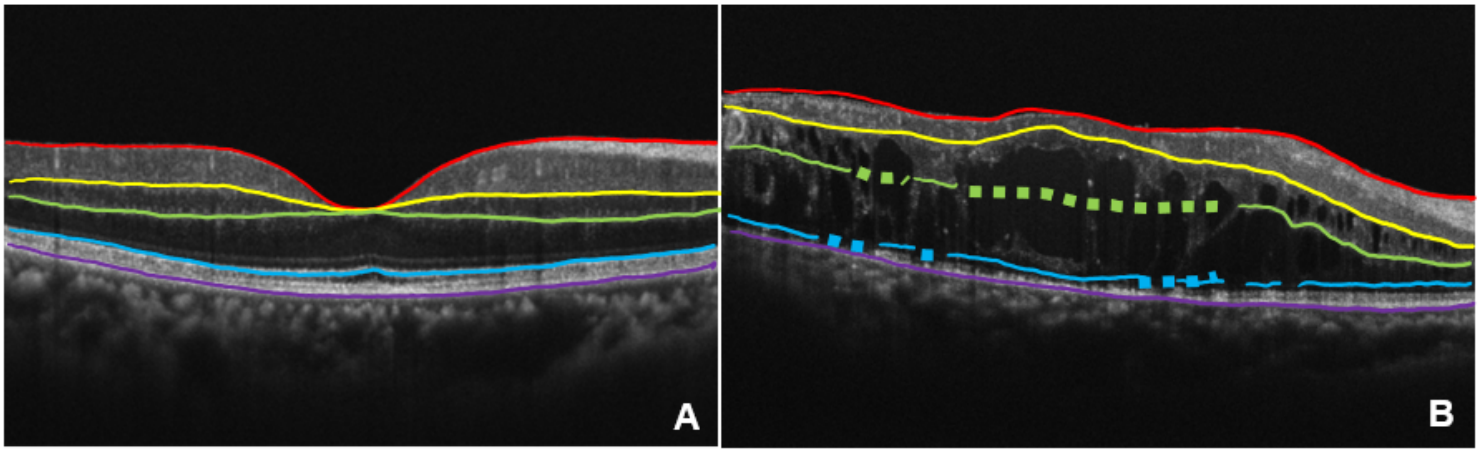

Figure 8

Example of segmentation using the Uncertainty Index. Normal eye (A) and diabetic retinopathy eye (B). The boundary line of the layer is drawn using AI in this study. As each boundary line is represented by a series of dots, there is an entropy value for each dot. The points where the value is high are delineated by a thick dotted line, and the points where the value is small are delineated by a thin solid line. Red: ILM, Yellow: IPL/INL, Green: OPL/ONL, Blue: EZ, Purple: RPE/BM

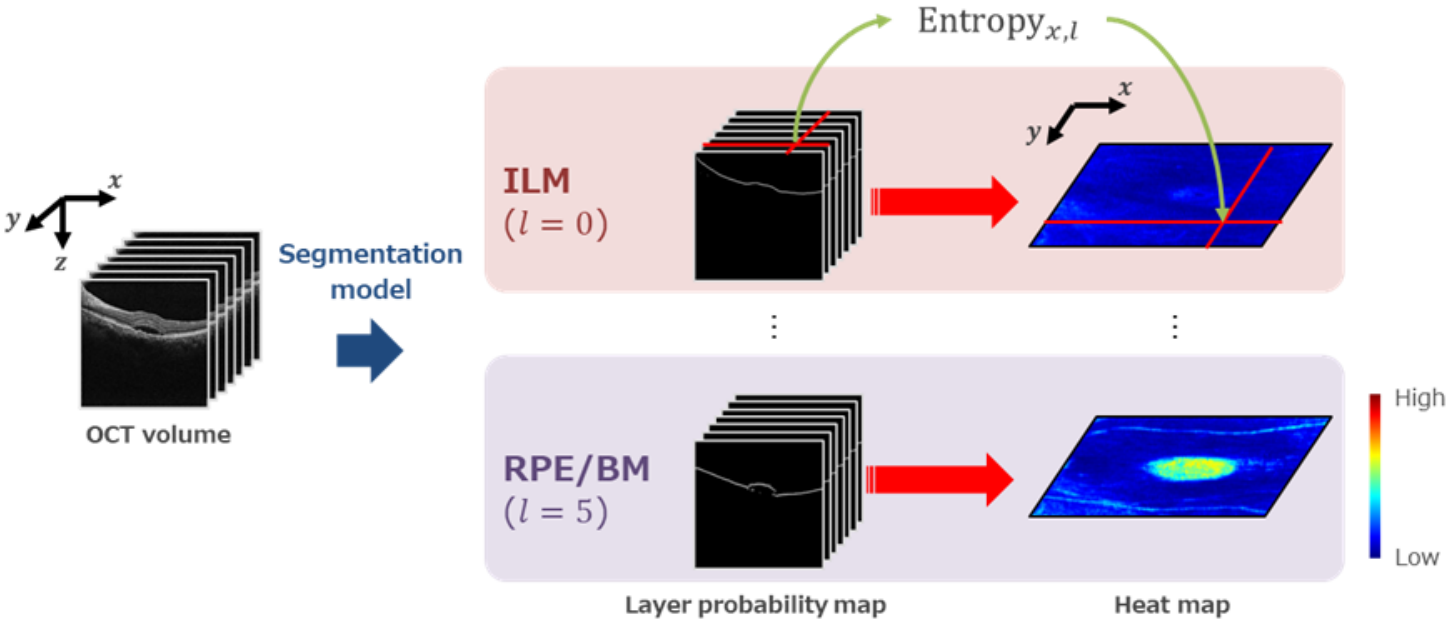

Figure 9

By calculating entropy in the volume scan, it is possible to create a heatmap for each layer of the retina and visualize the part with high entropy.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarydataSciRep.docx
- dataset.xlsx