



Published in final edited form as:

Br J Ophthalmol. 2022 September ; 106(9): 1301–1307. doi:10.1136/bjophthalmol-2020-317817.

Computer-aided detection (CADe) and abnormality score for the outer retinal layer in optical coherence tomography

Tyler Hyungtaek Rim, MD, MBA^{1,2,3,†}, Aaron Lee, MD, MSCI^{4,†}, Daniel Shu Wei Ting, MD^{1,2}, Kelvin Yi Chong Teo, MD^{1,2}, Henrik Hee Seung Yang¹ [MD candidate], Hyeonmin Kim⁵, Geunyoung Lee⁵, Zhen Ling Teo, MBBS, MRCS(Ed)¹, Alvin Wei Jun Teo, MBBS¹, Kengo Takahashi, MD⁶, Tae Keun Yoo, MD⁷, Seong Eun Kim, MD⁸, Yasuo Yanagi, MD, PhD⁶, Ching-Yu Cheng, MD, PhD^{1,2}, Sung Soo Kim, MD, PhD⁸, Tien Yin Wong, MD, PhD^{1,2}, Chui Ming Gemmy Cheung, MD^{1,2}

¹Singapore Eye Research Institute, Singapore National Eye Centre, Singapore

²Duke-NUS Medical School, National University of Singapore, Singapore

³Department of Ophthalmology, Yonsei University College of Medicine, Graduate School, Seoul, Korea

⁴Department of Ophthalmology, University of Washington School of Medicine, Seattle, Washington

⁵MediWhale, Seoul, South Korea

⁶Department of Ophthalmology, Asahikawa Medical University, Hokkaido, Japan

⁷Department of Ophthalmology, Aerospace Medical Center, Republic of Korea Air Force, Cheongju, South Korea

⁸Department of Ophthalmology, Severance Hospital, Institute of Vision Research, Yonsei University College of Medicine, Seoul, Korea

Corresponding author: Sung Soo Kim, MD, PhD. Department of Ophthalmology, Severance Hospital, Institute of Vision Research, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea, Tel: 82-10-2191-6313; Fax: 82-2-312-0541; semekim@yuhs.ac; **Gemmy Chui Ming Gemmy, MD, PhD.** Singapore Eye Research Institute, Singapore National Eye Centre, 11 Third Hospital Avenue, Singapore 168751, Phone: (65) 69881460, Fax: (65) 62263995; gemmy.cheung.c.m@snecc.com.sg.

[†]contributed equally

Author contributions statement

THR, AL, DSWT, C-YC, SSK, TYW, and CMGC conceptualized the study.

THR, AL, DSWT, KYCT, HHSY, HK, GL, ZLT, AWJT, KT, TKY, and SEK reviewed the literature.

THR, AL, DSWT, KYCT, C-YC, SSK, TYW, and CMGC designed the study.

THR, AL, KYCT, KT, TKY, SEK, YY, C-YC, SSK, TYW, and CMGC collected the data.

GL and HK developed the algorithm.

THR, AL, and C-YC analyzed data.

All authors contributed to data interpretation.

THR, AL, HHSY, ZLT, AWJT, TYW, and CMGC drafted the manuscript.

DSWT, KYCT, YY, C-YC, SSK, TYW, and CMGC did the critical revision.

All authors read and approved the final report.

Competing interests

THR was a scientific advisor to a start-up company called Medi-whale, Inc. He received stock as a part of the standard compensation package. DSWT and TYW hold patents on a deep learning system for the detection of retinal diseases and these patents are not directly related to this study. TYW has received consulting fees from Allergan, Bayer, Boehringer-Ingelheim, Genentech, Merck, Novartis, Oxurion, Roche, and Samsung Bioepis. TYW is a co-founder of Plano and EyRiS. DSWT is a co-founder of EyRiS. Potential conflicts of interests are managed according to institutional policies of the Singapore Health System (SingHealth) and the National University of Singapore. HK and GL are employee of Medi Whale Inc. All other authors declare no competing interests.

Abstract

Background: To develop computer-aided detection (CADE) of ORL abnormalities in the retinal pigmented epithelium (RPE), interdigitation zone (IZ), and ellipsoid zone (EZ) via optical coherence tomography (OCT).

Methods: In this retrospective study, healthy participants with normal ORL, and patients with abnormality of ORL including choroidal neovascularization (CNV) or retinitis pigmentosa (RP) were included. First, an automatic segmentation deep learning (DL) algorithm, CADE, was developed for the three outer retinal layers using 120 handcraft masks of ORL. This automatic segmentation algorithm generated 4000 segmentations, which included 2000 images with normal ORL and 2000 (1000 CNV and 1000 RP) images with focal or wide defects in ORL. Second, based on the automatically generated segmentation images, a binary classifier (normal versus abnormal) was developed. Results were evaluated by area under the receiver operating characteristic curve (AUC).

Results: The DL algorithm achieved an AUC of 0.987 (95% Confidence Interval [CI], 0.980-0.994) for individual image evaluation in the internal test set of 800 images. In addition, performance analysis of a publicly available external test set (n=968) had an AUC of 0.952 (95% CI, 0.938-0.966), and a second clinical external test set (n=1,146) had an AUC of 0.971 (95% CI, 0.962-0.980). Moreover, the CADE highlighted well normal parts of ORL and omitted highlights in abnormal ORLs of CNV and RP.

Conclusion: The CADE can use OCT images to segment ORL and differentiate between normal ORL and abnormal ORL. The CADE classifier also performs visualization and may aid future physician diagnosis and clinical applications.

Keywords

Deep learning; optical coherence tomography; computer-aided detection; outer retinal layer; retinal pigmented epithelium; interdigitation zone; ellipsoid zone; choroidal neovascularization; retinitis pigmentosa; drusen; diabetic macular edema

Introduction

Deep Learning (DL) is a novel artificial intelligence (AI) technology in which convoluted neural networks (CNNs) are programmed to optimize a specific performance criterion using large datasets with a known outcome.¹⁻³ Robust performance using these algorithms has been reported for outcomes as varied as identifying candidate patients for corneal refractive surgery,^{4 5} to classification of retinal diseases like diabetic retinopathy (DR) and age-related macular degeneration (AMD),⁶⁻⁹ as well as detection of glaucoma¹⁰ and retinopathy of prematurity.^{11 12}

Optical coherence tomography is a well-established diagnostic imaging technique for retinal disease. Correlations between pathophysiological markers seen in OCT images and visual function have been investigated in several retinal diseases.¹³⁻¹⁶ In particular, the photoreceptor layer integrity in OCT was found to robustly correlate with visual acuity compared to retinal thickness alone. As such, photoreceptor layer continuity as a predictive

indicator of visual acuity in various macular diseases have been studied.¹⁷ The currently agreed upon signs of photoreceptor damage or disruption in OCT are loss of integrity of the ellipsoid zone (EZ) and interdigitation zone (IZ) bands.¹⁸ Attenuation, discontinuity or disruption of these bands have been reported as likely hallmarks of photoreceptor dysfunction or damage in a variety of retinal diseases.^{17 19 20} In response to these new markers, OCT segmentation is quickly becoming a key evaluation tool.²¹ While existing research mainly focused on the pathogenesis and changes to subretinal fluid (SRF) in retinal diseases, there is a void of studies that examine ORL.

Considering the declining cost of OCT instruments and increasing accessibility, we believe that OCT will supplement general eye screening with retinal photography in the near future. Once OCTs become prevalent in the primary care setting, because the automated system helps healthcare professionals overtly visualize the ORL integrity, the operation is suitable for non-retinal specialists and novice ORL graders, with the greatest support advantage in cases of subtle ORL abnormality, especially in mass screening setting.

Therefore, to address this, we developed a DL algorithm that automatically analyzes ORL from OCT images. The algorithm is a multi-step build comprised of an automatic computer-aided detection (CADe) segmentation visualizer, which differentiates the retinal pigment epithelial (RPE) layer, EZ, and IZ, and a binary classifier, which further sorts the dataset into normal vs abnormal ORLs. We hypothesize that this automated OCT CADe will ultimately make primary screening that currently depend only on retinal photography more efficient in the near future.

Methods

Study Design

This was a retrospective cross-sectional study. The study was approved by the local ethical committee of Yonsei University (4-2019-0442). The ethics committee waived written informed consent in view of image and clinical data de-identification. This study adhered to the tenets of the Declaration of Helsinki.

Participants and Image database

For algorithm development, we obtained OCT images and clinical data of all participants at Severance Hospital, Yonsei University between 2005 and 2017.²² We extracted patients' clinical information from electronic medical records and the physician's order communication system that was stored in the clinical data repository and data warehousing in Severance Hospital.²³ Optical coherence tomography images included 5 images centered on the fovea per participant. Original Images were extracted from the database in PNG (Portable Network Graphics) format then underwent noise removal and contrast enhancement.²² Initial extraction yielded images of two sizes; 768 X 496 (9 mm scan) and 512 X 496 (6 mm scan). For balanced processing, we transformed 512 X 496 images to 768 X 496 size by assigning a blank area outside of the scanned area.

Normal Outer Retinal Layer

Based on the international panel of OCT experts consensus in Figure 1, a normal ORL includes three different layers - 1) RPE layer including the RPE and Bruch's membrane, 2) the IZ, and 3) the EZ band.^{18 24 25} To be qualified as a normal ORL, all three layers must be absent of any abnormality, and pathologic elevation or depression. Normal ORL OCT images were manually selected by a retinal specialist (T.H.R.) from the pooled OCT images, which were extracted from those who underwent OCT at first visit without further OCT examinations during the follow-up period in ophthalmology clinics up to 2017.

Abnormal Outer Retinal Layer

To construct the representative abnormal ORL dataset, we used two representative retinal diseases; CNV and pigmented retinitis (RP). The information on the preliminary database of two retinal diseases of CNV²⁶ and RP²⁷ was provided in Supplementary Document 1. Abnormal ORL was defined as any absence, attenuation, discontinuity, disruption, elevation or depression of the RPE, EZ, or IZ layer within OCT image.^{17 19 20} We selected OCT images with abnormal ORL manually (T.H.R.) from CNV database and RP database because some images did not include ORL abnormality. We manually excluded normal ORLs with underlying retinal disease and poor quality of images.

Development of segmentation model, Model 1

The overall DL algorithm development process is provided in Figure 1. We developed **Model 1** for segmentation of normal ORL. 120 normal OCT images were used as representatives to manually set the ground truth of the region of interest (ROI). We used hand-crafted 100 ROI masks for the training set and 20 ROI masks for internal test set. To reduce overfitting, data augmentation techniques were applied. Each and every training sample image underwent the following transformations: random flip left-right, random rotation from -20 to 20 degree angle, random brightness enhancement with max delta 0.5, and Gaussian Blur with kernel size 3. We constructed a CNN based on U-Net,²⁸ and every CNN block in U-Net was followed by batch normalizations for regularization of the model. Adam Optimizer with fixed weight decay and a learning rate of 1e-5 during 140 epochs was used to train **Model 1**.

Development of classification model, Model 2

Model 2 was developed to further classify OCT images into normal and abnormal ORL (Figure 1). 4000 segmentation masks (2000 normal vs 1000 CNV and 1000 RP) were generated from DL **Model 1** (3200 masks for training set and 800 masks for test set). Given the binary classification, data augmentation techniques were not applied, and VGG16 was used. Adam Optimizer with fixed weight decay and a learning rate of 1e-4 during 100 epochs was used to train **Model 2**. The abnormality score, a representation of the probability prediction score that ranged from zero to one, was outputted by the algorithm. The abnormality score of zero indicated a normal ORL, and the abnormality score of one indicated an abnormal ORL. A binary screening cutoff value was necessary, and thus determined based on the highest Youden Index (the highest sensitivity + specificity). Results were color-coded with indicators; a green circle is normal and a red circle is abnormal.

External test sets

Two different external test sets were prepared to assess **Model 2** performance. The first set is a publicly available OCT database²⁹ consisting of four categories; normal, drusen, CNV, and DME. Each category contained 242 B-scans for a total of 968 images. Since this is a publicly available dataset, no further manual selection was done. The second set is comprised of clinical data from the Singapore Epidemiology of Eye Diseases (SEED) study,³⁰ and Asahikawa University, Japan. We selected OCT images with abnormal ORL manually from the mixed retinal disease [drusen, CNV, Geographic atrophy, macular hole, and laser scar], CNV, and RP, because some images did not have ORL abnormality. These clinical data included 232 images of mixed retinal disease from SEED and 370 images of CNV and 302 images of RP from Asahikawa University.

Statistical Analysis

TensorFlow (<http://tensorflow.org>), an open-source software library for machine intelligence, was used in the training and evaluation of the models. Tools including NumPy, SciPy, matplotlib, scikit-learn was used to process the data and analyze the Receiver Operating Characteristic (ROC) curve.³¹ We estimated intersection of union (IoU), accuracy, and boundary F1 (BF) score for segmentation, and receiver operating characteristic (ROC) curves, and area under the ROC curves for classification. The output performance per each individual image was evaluated with a 95% Confidential Interval (CI). Sensitivity and specificity were determined from the Youden Index.

Results

The performance of **Model 1** is shown in Supplementary Table 1. In RPE, IZ, and EZ, IoUs were 0.83 (0.80-0.86), 0.78 (0.74-0.82), and 0.85 (0.83-0.86), respectively. For binary classification of normal versus abnormal ORL (**Model 2**), we achieved good performance with AUC of 0.978 (95% CI, 0.980-0.994) at the image level in internal testing. Sensitivity and specificity with optimal cutoffs based on the highest Youden Index was 97.8% and 95.1%. Supplementary Figure 1 provides the ROC curve in internal test set.

Based on this binary classification, the distribution of the abnormality score is provided in Supplementary Figure 2. Mean abnormality score was 0.080 (standard deviation [SD], 0.208) in normal ORL group and 0.963 (SD, 0.142) in abnormal ORL group (0.937 [SD, 0.192] for CNV and 0.982 [SD, 0.085] for retinitis pigmentosa), and the optimal cutoff value was determined as 0.658 based on Youden method. Overall, the distribution of abnormality scores for RP is closer to 1.00 (100% abnormality) than that for CNV. This distribution is consistent with the higher performance in retinitis pigmentosa than in CNV shown in Supplementary Figure 1.

In the publicly available external test set of 986 OCT images (242 normal, 242, drusen, 242 CVN, and 242 DME images), binary classifier for normal vs abnormal ORL demonstrated good performance with AUC of 0.952 (95% CI, 0.938-0.966) at the image level (Figure 2A, 1st curve). In disease-specific performance, AUCs for drusen, CNV, and DME were 0.959 (95% CI, 0.943-0.976), 0.991 (95% CI, 0.984-0.997), and 0.906 (95% CI, 0.878-0.934),

respectively (Figure 2A, 2nd to 4th curves). In the clinical external test set of 1,446 images (242 normal ORL, 232 abnormal ORL from mixed disease, and 370, and 302 abnormal ORL from CNV, and RP, respectively), AUC was 0.971 (95% CI, 0.962-0.980) for normal vs abnormal ORL (Figure 2B, 1st curve). The performance from the SEED data showed AUC of 0.918 (95% CI, 0.891-0.945) for any abnormal ORL from mixed retinal disease. In the clinical data from Asahikawa University, the AUCs were 0.987 (95% CI, 0.980-0.995) in abnormal ORL images with CNV, and 0.991 (95% CI, 0.986-0.996) in abnormal ORL images with RP.

Representative images of CADe with an abnormality score is provided in Figure 3. Figure 3A shows a normal ORL with an abnormality score of 0%, a green circle indicator, and well demarcated, continuous predicted segmentation highlights for all three layers of a normal ORL. Figure 3B shows a severe RP case with an abnormality score of 100%, a red circle indicator, and CADe segmentation with sparse, staccato highlights that represent the intact areas of EZ and IZ. Meanwhile, the mild RP case in Figure 3C with an abnormality score of 0.69 shows a relatively small area of absent EZ and IZ layers along the right corner. Supplementary Figure 3 provides more cases akin to Figure 3C with relatively minimal abnormal ORL in RP OCT images with an abnormality score of <100%. Figure 4 shows the CNV images in the internal test set and Figure 5 shows the CADe on external test set.

Discussion

Our study demonstrates a system for assessing the integrity of the ORL and detecting ORL abnormality in a number of clinical conditions. The output results of CADe utilize highlighted visualizations for each ORL, which may support rapid screening, especially in settings not supported by clinical experts. Furthermore, our binary classifier demonstrated good performance in sorting abnormal ORL and normal ORL with an AUC of 0.952 in the publicly available external data set, and 0.971 in the clinical data set from two different institutions.

As prevalent and routine as OCT images have become, there is a conjunct necessity to efficiently assess retinal segmentation in order to decipher and extract the most clinical utility.²¹ Prior studies designed automated algorithms that detect or quantify intraretinal and subretinal fluid in OCT images of CNV, diabetic macular edema, and retinal vein occlusion with high accuracy.³²⁻³⁴ The clinician may potentially work with state of the art AI with real-time fluid detection and quantification in a not so distant time. A combined system based on 3-dimensional tissue-segmentation map and classification was previously developed for providing one of four referral suggestions (e.g. urgent) and diagnosis.³⁵ However, complex algorithm is more likely to cause errors with limited generalization in a real-world setting.³⁶ Therefore, simple approach is not necessarily a disadvantage because it is easy for the user to understand and the errors can be reduced. An additional feature that prospective AI-aided clinician support systems will require for OCT image analysis is assessment of ORL integrity, a crucial biomarker for vision.^{17 19 20} ORL integrity may serve as an intuitive biomarker of visual function that is applicable to most retinal diseases and a more discernable anatomical structure than the intraretinal cyst or hyperreflective material.

A two tiered system build of DL-based segmentation and classification was necessary to streamline output data into a clinically approachable form for healthcare professionals. The user friendly interface and result dispensation of clinical support systems are often overlooked. One such DL-system visualization technique is the heatmap.²² Heatmaps show how the machine learns the classification and how the algorithm determines important features for the classification. However, while some heatmaps feature clear pathophysiological structures, others highlight structures of no known clinical correlation (e.g. black parts in OCT behind the choroid) that render physician interpretation moot.²² Therefore, a pathophysiological approach grounded in clinically relevant anatomical structures of interest, such as the ORL, is necessary. Our system attempts to deliver on these points by clearly visualized, automated ORL segmentation and screening of abnormalities.

It is important to differentiate that the binary classifier (**Model 2**) is not a classifier for retinal diseases but a classifier for ORL abnormalities as not all retinal diseases present with ORL abnormalities. The internal test set was manually selected only when there was an ORL abnormality, and the performance was good with AUC of 0.987. The external test set included retinal diseases but some images, specifically the DME images, do not have ORL abnormalities, as shown in Figure 5D. Correspondingly, the DME cases had the lowest performance (Figure 2A).

Because the CADe system helps healthcare professionals overtly visualize the ORL integrity, the operation is suitable for non-retinal specialists and novice ORL graders, with the greatest support advantage in cases of subtle ORL abnormality such as mild RP cases described in Supplementary Figure 3, which contain slight aberrations that are difficult to pick up to the untrained eye. Moreover, considering the declining cost of OCT instruments and increasing accessibility, we believe that OCT will supplement general eye screening with retinal photography in the near future. Once OCTs become prevalent in the primary care setting, CADe will embolden the general practitioner.

The abnormality score, which is the binary classification probability prediction score in **Model 2**, requires further clarification. This score is an indicator of ORL normality. As such, the score no linear relation to abnormal length of ORL. For example, abnormality score of 0.8 does not translate to damage or omission of 80% of the ORL on OCT scan. We stress that the abnormality score should be interpreted as the probability of having an abnormal ORL compared to a normal ORL. The abnormality score is a broadly accepted unit in medical artificial intelligence applications including detection of lung nodules in chest X-rays,³⁷ breast cancer in mammography,³⁸ and DL-based triage score generation in the emergency department.³⁹ Additionally, further verification must be done to ascertain how the abnormality score translates in the clinical setting. For example, this score may be associated with treatment response after intravitreal injection in AMD.

Our study has several limitations. **First**, our segmentation algorithm was developed upon ground truth defined by only normal ORLs and not from abnormal ORLs. As such, the accuracy is limited in abnormal ORLs. Accurate ORL segmentation for all ORLs requires further development of **model 1** with the correct ground truth for abnormal ORL. However, our CADe was shown to be applicable to the abnormal ORL even though **model 1** was

developed only on normal ORLs. Moreover, the performance of the binary classifier, which was developed based on these segmented masks demonstrated good performance in screening abnormal ORLs. **Second**, our DL algorithm was trained on a very specific set of images with extremes in phenotypes. The performance of this model may be degraded in the real-world setting when other ocular pathology confounders are present. **Third**, this model may not apply to images acquired from different OCT platforms. However, performing an extensive external test is beyond our scope, and the primary purpose of this study was to develop a CADe that visualized and labeled intact ORL as a proof of concept clinical decision support system. Future studies would include images from different OCT manufacturers for further generalizability.

In conclusion, we developed a DL algorithm capable of automated segmentation for normal ORLs. The CADe successfully highlighted normal ORLs when we applied it to abnormal ORLs. Furthermore, we developed a DL-based ORL classifier with a computed abnormality score that demonstrated good performance in both public and clinical external test sets. As AI and DL permeates healthcare, we envisage a CADe system that aids inexperienced healthcare professionals as well as retinal specialists in OCT image interpretation and patient care delivery.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This study was funded by the National Medical Research Council of Singapore (NMRC/OFLCG/004a/2018; NMRC/CIRG/1488/2018) and of National Eye Institute of the United States (K23EY029246).

Data availability statement

Data cannot be shared publicly due to the violation of patient privacy and lack of informed consent for data sharing. Data are available from the Yonsei University, Department of Ophthalmology (contact Prof. Sung Soo KIM, semekim@yuhs.ac) for researchers who meet the criteria for access to confidential data.

References

1. Ting DSW, Peng L, Varadarajan AV, et al. Deep learning in ophthalmology: The technical and clinical considerations. *Prog Retin Eye Res* 2019.
2. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *British Journal of Ophthalmology* 2019;103:167–75. [PubMed: 30361278]
3. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, et al. Artificial intelligence in retina. *Prog Retin Eye Res* 2018;67:1–29. [PubMed: 30076935]
4. Yoo TK, Ryu IH, Lee G, et al. Adopting machine learning to automatically identify candidate patients for corneal refractive surgery. *npj Digital Medicine* 2019;2:59. [PubMed: 31304405]
5. Yoo TK, Ryu IH, Choi H, et al. Explainable Machine Learning Approach as a Tool to Understand Factors Used to Select the Refractive Surgery Technique on the Expert Level. *Translational Vision Science & Technology* 2020;9:8-.

6. Grassmann F, Mengelkamp J, Brandl C, 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* 2018;125:1410–20. [PubMed: 29653860]
7. Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA ophthalmology* 2017;135:1170–6. [PubMed: 28973096]
8. Burlina PM, Joshi N, Pacheco KD, 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* 2018;136:1359–66. [PubMed: 30242349]
9. Ting DSW, Cheung CY-L, Lim G, 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* 2017;318:2211–23. [PubMed: 29234807]
10. Hood DC, De Moraes CG. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology* 2018;125:1207–8. [PubMed: 30032794]
11. Brown JM, Campbell JP, Beers A, et al. Automated Diagnosis of Plus Disease in Retinopathy of Prematurity Using Deep Convolutional Neural Networks. *JAMA ophthalmology* 2018;136:803. [PubMed: 29801159]
12. Ting DSW, Wu W-C, Toth C. Deep learning for retinopathy of prematurity screening. *British Journal of Ophthalmology* 2018;bjophthalmol-2018-313290.
13. Keane PA, Liakopoulos S, Chang KT, et al. Relationship between optical coherence tomography retinal parameters and visual acuity in neovascular age-related macular degeneration. *Ophthalmology* 2008;115:2206–14. [PubMed: 18930551]
14. Alasil T, Keane PA, Updike JF, et al. Relationship between optical coherence tomography retinal parameters and visual acuity in diabetic macular edema. *Ophthalmology* 2010;117:2379–86. [PubMed: 20561684]
15. Hayashi H, Yamashiro K, Tsujikawa A, et al. Association between foveal photoreceptor integrity and visual outcome in neovascular age-related macular degeneration. *American journal of ophthalmology* 2009;148:83–9. e1. [PubMed: 19327745]
16. Forooghian F, Stetson PF, Meyer SA, et al. Relationship between photoreceptor outer segment length and visual acuity in diabetic macular edema. *Retina (Philadelphia, Pa.)* 2010;30:63.
17. Schliesser JA, Gallimore G, Kunjukunju N, et al. Clinical application of optical coherence tomography in combination with functional diagnostics: advantages and limitations for diagnosis and assessment of therapy outcome in central serous chorioretinopathy. *Clin Ophthalmol* 2014;8:2337–45. [PubMed: 25473259]
18. Spaide RF, Curcio CA. Anatomical correlates to the bands seen in the outer retina by optical coherence tomography: literature review and model. *Retina (Philadelphia, Pa.)* 2011;31:1609.
19. Hagiwara A, Mitamura Y, Kumagai K, et al. Photoreceptor impairment on optical coherence tomographic images in patients with retinitis pigmentosa. *Br J Ophthalmol* 2013;97:237–8. [PubMed: 23172877]
20. Aizawa S, Mitamura Y, Hagiwara A, et al. Changes of fundus autofluorescence, photoreceptor inner and outer segment junction line, and visual function in patients with retinitis pigmentosa. *Clin Exp Ophthalmol* 2010;38:597–604. [PubMed: 20456441]
21. DeBuc DC. A review of algorithms for segmentation of retinal image data using optical coherence tomography. *Image Segmentation* 2011;1:15–54.
22. Lee CS, Baughman DM, Lee AY. Deep learning is effective for the classification of OCT images of normal versus Age-related Macular Degeneration. *Ophthalmol Retina* 2017;1:322–7. [PubMed: 30693348]
23. Chang BC, Kim NH, Kim YA, et al. Ubiquitous-severance hospital project: implementation and results. *Healthc Inform Res* 2010;16:60–4. [PubMed: 21818425]
24. Staurengi G, Sadda S, Chakravarthy U, et al. Proposed lexicon for anatomic landmarks in normal posterior segment spectral-domain optical coherence tomography: the IN•OCT consensus. *Ophthalmology* 2014;121:1572–8. [PubMed: 24755005]

25. Saxena S, Srivastav K, Cheung CM, et al. Photoreceptor inner segment ellipsoid band integrity on spectral domain optical coherence tomography. *Clin Ophthalmol* 2014;8:2507–22. [PubMed: 25525329]
26. Mitchell P, Liew G, Gopinath B, et al. Age-related macular degeneration. *Lancet* 2018;392:1147–59. [PubMed: 30303083]
27. Rim TH, Park HW, Kim DW, et al. Four-year nationwide incidence of retinitis pigmentosa in South Korea: a population-based retrospective study from 2011 to 2014. *BMJ Open* 2017;7:e015531.
28. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*: Springer, 2015:234–41.
29. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 2018;172:1122–31 e9. [PubMed: 29474911]
30. Tan GS, Gan A, Sabanayagam C, et al. Ethnic Differences in the Prevalence and Risk Factors of Diabetic Retinopathy: The Singapore Epidemiology of Eye Diseases Study. *Ophthalmology* 2018;125:529–36. [PubMed: 29217148]
31. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 2011;12:2825–30.
32. Chakravarthy U, Goldenberg D, Young G, et al. Automated Identification of Lesion Activity in Neovascular Age-Related Macular Degeneration. *Ophthalmology* 2016;123:1731–6. [PubMed: 27206840]
33. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning. *Ophthalmology* 2018;125:549–58. [PubMed: 29224926]
34. Lee H, Kang KE, Chung H, et al. Automated Segmentation of Lesions Including Subretinal Hyperreflective Material in Neovascular Age-related Macular Degeneration. *Am J Ophthalmol* 2018;191:64–75. [PubMed: 29655643]
35. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24:1342–50. [PubMed: 30104768]
36. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019;17:195. [PubMed: 31665002]
37. Nam JG, Park S, Hwang EJ, et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* 2019;290:218–28. [PubMed: 30251934]
38. Kim EK, Kim HE, Han K, et al. Applying Data-driven Imaging Biomarker in Mammography for Breast Cancer Screening: Preliminary Study. *Sci Rep* 2018;8:2762. [PubMed: 29426948]
39. Kwon J-M, Lee Y, Lee Y, et al. Validation of deep-learning-based triage and acuity score using a large national dataset. *PloS one* 2018;13:e0205836. [PubMed: 30321231]

Synopsis

We aimed to develop computer-aided detection (CADe) of outer retinal layer (ORL) abnormalities via optical coherence tomography (OCT). The results of CAdE utilize highlighted visualizations for ORL and our binary classifier demonstrated excellent performance.

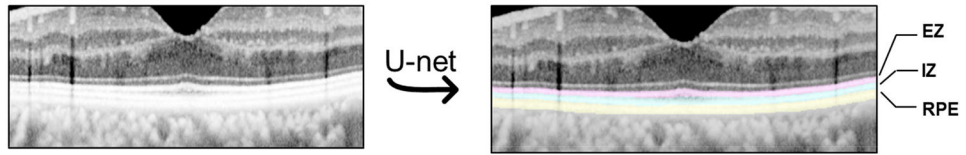
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

A Model 1 Segmentation



B Model 2 Binary classification

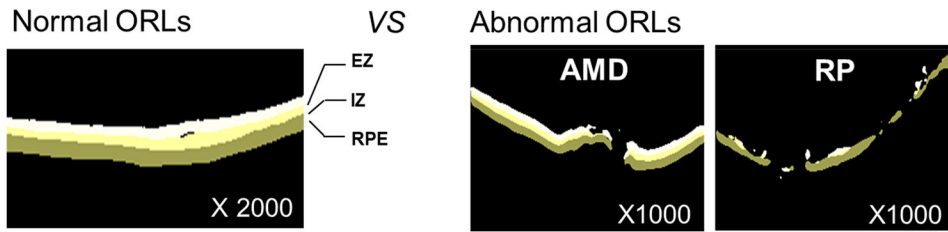
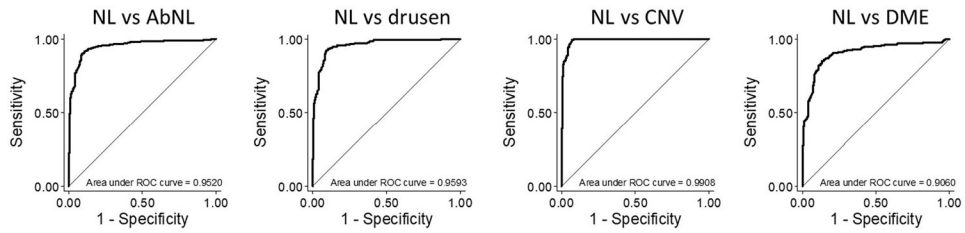


Figure 1. Outer retina layer (ORL) and handcraft mask and Study design

EZ = ellipsoid zone; IZ = interdigitation zone; RPE=retinal pigment epithelium and Bruch's membrane.

A. Publicly available external test set



B. Clinical external test set

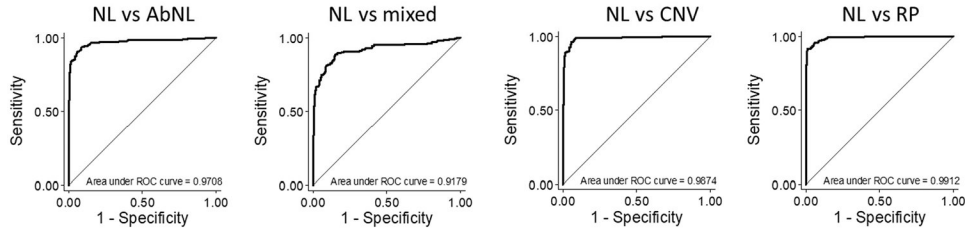


Figure 2. ROC curves for external test sets.

AbNL=abnormal; CNV=choroidal neovascularization; DME=diabetic retinopathy;

NL=normal; RP=retinitis pigmentosa.

Mixed = abnormal outer retinal layer from mixed retinal disease.

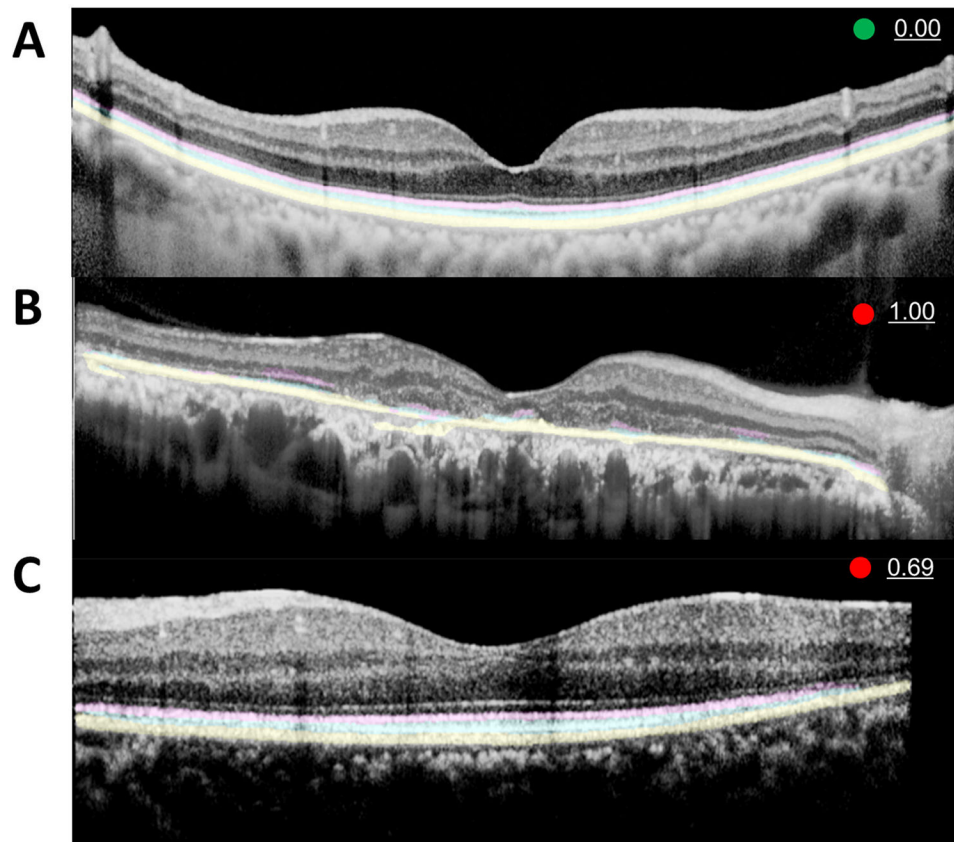


Figure 3. Representative examples of computer aided detection system.

(A) Computer aided detection (CAdE) system highlights of the normal outer retina including ellipsoid zone (EZ, pink), interdigitation zone (IZ, sky blue), and retinal pigment epithelium and Bruch's membrane (RPE, yellow) are clearly demarcated with zero abnormality score and green circle. (B) A severe case of widespread disruption of EZ and IZ in an OCT image of retinitis pigmentosa (RP) with 1.00 (100%) abnormality score. CAdE well-highlighted the remaining EZ and IZ. This will aid the physician in recognizing ORL disruption. (C) In a mild case of RP, the EZ and IZ disruption is limited to the right corner, which is discernable on CAdE rendered ORL with an intact RPE layer with prematurely tapered pink and blue highlights. The abnormality score of 0.69 is marginally higher than our cutoff value of 0.66 (see results section)

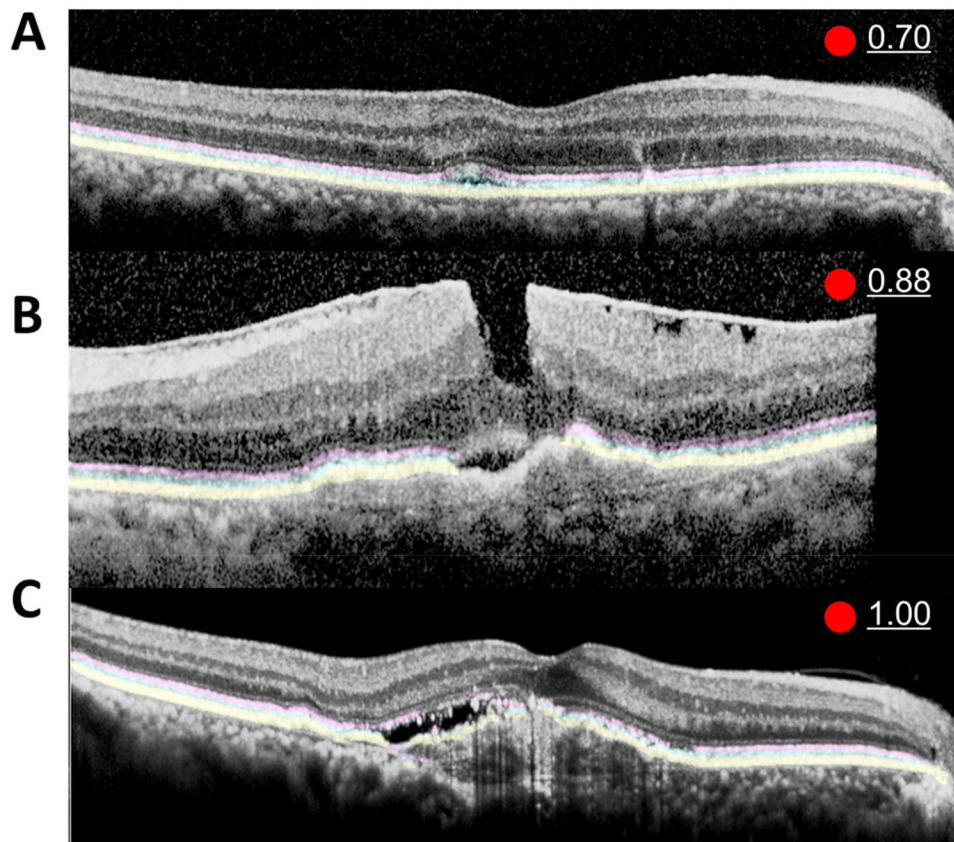


Figure 4. Representative examples of computer aided detection system for choroidal neovascularization.

Computer aided detection (CAdE) system clearly highlighted the normal outer retina including ellipsoid zone (EZ, pink), interdigitation zone (IZ, sky blue), and retinal pigment epithelium and Bruch's membrane (RPE, yellow) in the intact ORL regions of the OCT image from a patient with age-related macular degeneration (AMD). There is small amount of subretinal fluid (A and B). Partial disruption of the pink EZ is noted in the mild case (A), and both pink and sky-blue EZ and IZ are disrupted in more severe cases (B). The mild case of AMD (A) has a lower abnormality score of 0.70 compared to the 0.88 score of the severe case of AMD (B). There is a thick epiretinal membrane in (B) that CAdE incorrectly did not recognize as an ORL. In contrast, the disruption of IZ on the left half and normal IZ on the right half was correctly denoted by CAdE and a 1.00 abnormality score was reported in severe AMD case (C).

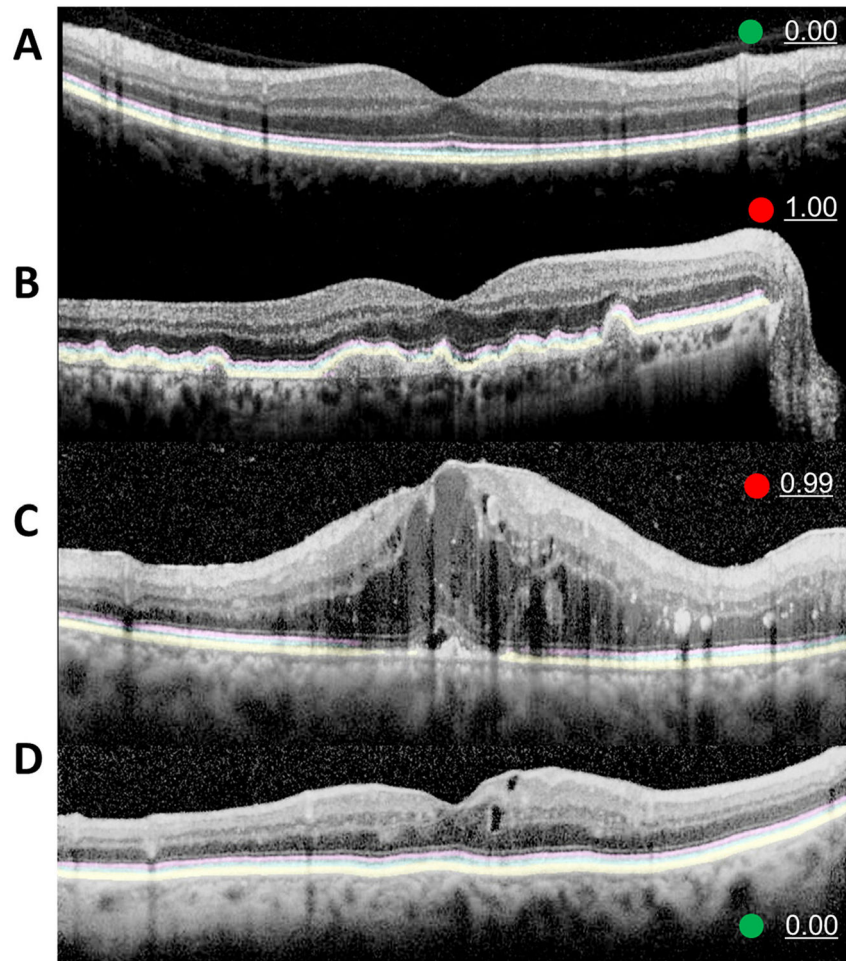


Figure 5. Representative examples of computer aided detection system for publicly available external test set.

Computer aided detection (CADe) system correctly discerned with well-defined highlights in the normal outer retina (A) of an intact ORL in external test set. Our segmentation algorithm was developed by training only on normal ORLs, but it was capable of segmenting drusen with an intact but curved ORL (B). There is a small amount of subretinal fluid and shadowing due to intraretinal cyst in the case of the diabetic retinopathy (C) and the corresponding area has weakened highlight signals. Additionally, the abnormality score was zero (0.00) in a DME with a very small amount of intraretinal cyst without ORL abnormality (D). Because of such cases, the performance of the binary classifier for DME is rather low (results, Figure 2).