Thieme

Deep Learning-Based Retinal Layer Segmentation in Optical Coherence Tomography Scans of Patients with Inherited Retinal Diseases

Deep-Learning-basierte Segmentierung von Aufnahmen aus optischer Kohärenztomografie bei Patienten mit erblichen Netzhauterkrankungen

OPEN ACCESS

Authors

Franziska Eckardt[1], Robin Mittas[2], Nastassya Horlava[2], Johannes Schiefelbein[1], Ben Asani[1], Stylianos Michalakis[1], Maximilian Gerhardt[1], Claudia Priglinger[1], Daniel Keeser[3], Nikolaos Koutouleris[3], Siegfried Priglinger[1], Fabian Theis[2], Tingying Peng[2], Benedikt Schworm[1]

Affiliations

1 Department of Ophthalmology, LMU University Hospital, LMU Munich, Munich, Germany
2 Institute for Computational Biology, Helmholtz Munich, Munich, Germany
3 Department of Psychiatry und Psychotherapy, LMU University Hospital, LMU Munich, Munich, Germany

Key words

retina, inherited retinal disease, OCT

Schlüsselwörter

Retina, erbliche Netzhauterkrankungen, OCT

received	10.10.2023
accepted	23.10.2023
published online	

Bibliography

Klin Monatsbl Augenheilkd 2024
DOI 10.1055/a-2227-3742
ISSN 0023-2165

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence

Dr. Franziska Eckardt
Augenklinik und Poliklinik, LMU Klinikum München
Mathildenstraße 8, 80336 München, Germany
Phone: + 49 (0) 8 94 40 05 37 70, Fax: + 49 (0) 8 94 40 05 51 60
franziska.eckardt@med.uni-muenchen.de

ABSTRACT

Background In optical coherence tomography (OCT) scans of patients with inherited retinal diseases (IRDs), the measurement of the thickness of the outer nuclear layer (ONL) has been well established as a surrogate marker for photoreceptor preservation. Current automatic segmentation tools fail in OCT segmentation in IRDs, and manual segmentation is time-consuming.

Methods and Material Patients with IRD and an available OCT scan were screened for the present study. Additionally, OCT scans of patients without retinal disease were included to provide training data for artificial intelligence (AI). We trained a U-net-based model on healthy patients and applied a domain adaption technique to the IRD patients' scans.

Results We established an AI-based image segmentation algorithm that reliably segments the ONL in OCT scans of IRD patients. In a test dataset, the dice score of the algorithm was 98.7%. Furthermore, we generated thickness maps of the full retinal thickness and the ONL layer for each patient.

Conclusion Accurate segmentation of anatomical layers on OCT scans plays a crucial role for predictive models linking retinal structure to visual function. Our algorithm for segmentation of OCT images could provide the basis for further studies on IRDs.

ZUSAMMENFASSUNG

Hintergrund Bei der optischen Kohärenztomografie (OCT) von Patienten mit erblichen Netzhauterkrankungen hat sich die Messung der äußeren Körnerschichtdicke (ONL) als Marker für den Erhalt der Photorezeptoren bewährt. Derzeitige automatische Segmentierungsprogramme versagen bei der OCT-Segmentierung dieser Patienten, und die manuelle Segmentierung ist zeitaufwendig.

Methoden und Material Für die vorliegende Studie wurden Patienten mit erblichen Netzhauterkrankungen und der Ver-

fügbarkeit eines OCT-Scans eingeschlossen. Zusätzlich wurden OCT-Scans von Patienten ohne Netzhauterkrankung einbezogen, um Trainingsdaten für die künstliche Intelligenz (KI) zu generieren. Wir trainierten ein auf einem U-Netz basierendes Modell an gesunden Patienten und wendeten eine Anpassungsmethode auf die pathologisch veränderten Scans von Patienten an.

Ergebnisse Es wurde ein KI-basierter Bildsegmentierungsalgorithmus entwickelt, der die ONL in OCT-Scans von Patienten mit erblichen Netzhauterkrankungen zuverlässig segmentieren kann. In einem Testdatensatz lag der Dice-Score des Algorithmus bei 98,7 %. Außerdem erstellten wir für jeden Patienten Dickenkarten der gesamten Netzhautdicke und der ONL-Schicht.

Schlussfolgerung Eine präzise Segmentierung anatomischer Schichten auf OCT-Scans ist entscheidend für Prognosemodelle, die Netzhautstruktur und Sehfunktion korrelieren. Der hier vorgestellte OCT-Bildsegmentierungsalgorithmus könnte die Grundlage für weitere Studien bez. erblicher Netzhauterkrankungen darstellen.

Introduction

Inherited retinal diseases (IRDs) are a diverse group of genetic disorders that lead to progressive degeneration of the retina. These diseases encompass a wide range of conditions such as retinitis pigmentosa, Leber's congenital amaurosis, Stargardt disease, and many others. IRDs often lead to significant vision impairment and potentially blindness. In Europe, IRDs affect about 1 : 3000 individuals. Optical coherence tomography (OCT), as a noninvasive imaging test, has become an indispensable tool for diagnosing and monitoring the progression of retinal conditions for IRDs [1, 2]. OCT imaging allows an analysis and measurement of the retina's distinct layers, which can be crucial for confirming a diagnosis, guiding treatment plans, and assessing a response to therapies in the management of IRDs.

With the advent of artificial intelligence (AI) and machine learning, there have been several studies aiming to help disease diagnosis based on OCT. However, compared to more common ophthalmological diseases like diabetic retinopathy (DR) [3], age-related macular degeneration (AMD) [4], and glaucoma [5], AI studies specialized on IRDs are still scarce. In this work, we initiated a study for a deep learning-based retinal layer segmentation for IRD patients with a special focus on the outer nuclear layer (ONL). In IRDs, ONL thickness measurement has been well established as a surrogate marker for photoreceptor preservation. The goal of this work was to establish an AI-based image segmentation algorithm that reliably segments the ONL in OCT scans and provides a full retina and ONL thickness map of IRD patients.

Methods and Material

Participants and imaging

The clinical research database of the Department of Ophthalmology was screened for patients with a confirmed IRD diagnosis and available OCT scans. Normal data was collected from healthy participants without retinal disease. From every individual, both eyes were included in this study. Spectral-domain OCT and near infrared (NIR) confocal scanning-laser ophthalmoscopy was performed using the Spectralis HRA + OCT platform (Heidelberg Engineering GmbH, Heidelberg, Germany). Each OCT scan consisted of either 49 or 97 slides [one slide refers to one 2D image (B-scan) with a resolution of 496 × 512 pixels] and covers 20 × 20 degrees of the posterior pole centered to the fovea centralis. OCT scans were only included with a quality index of 20 or better and no blinking artifacts. OCT scans were exported as. dcm files from the manufacturer's software Heidelberg Eye Explorer (Version 1.10.4.0, Heidelberg Engineering GmbH, Heidelberg, Germany). The local ethics committee of the medical faculty gave ethical approval for this work (identifier 23–0392). The study adhered to the tenets of the Declaration of Helsinki.

Manually annotated OCT segmentation for training datasets

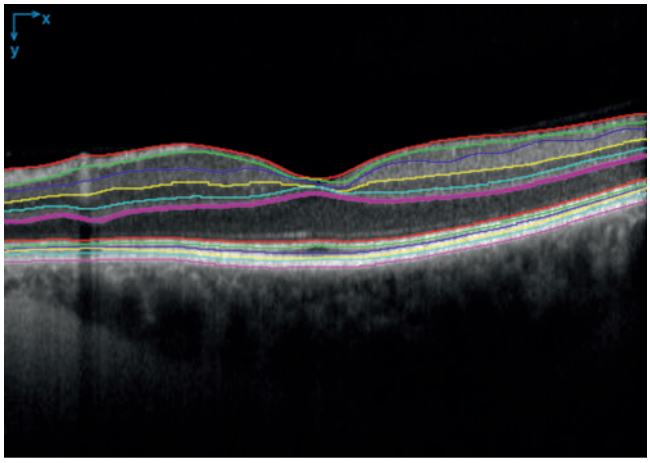
The annotations for the healthy retinal dataset were obtained from the software "OCTExplorer," which was developed by the IOWA University [6 – 8] and is free to use by academics. We used this software as its segmentation results on healthy retinal OCTs are reliable and only a small proportion of slides needed to be re-annotated by an expert. Furthermore, it is also possible to access the segmented files from the software to further process and transform the segmentation by a small helper function.

However, in IRD patients, the existing automatic segmentation tools, including the abovementioned "OCTExplorer," fail to reliably segment these scans due to structural alterations caused by the degeneration of retinal layers. Examples of OCTExplorer segmentations are shown in ▶ **Fig. 1** and **Fig. 2**.

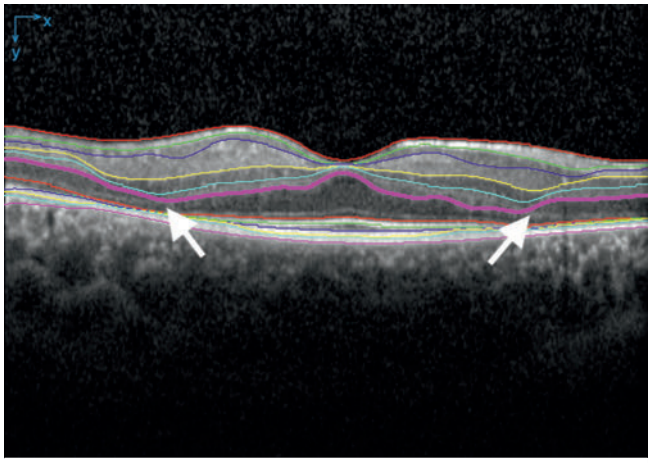
Consequently, we had to rely on expert annotations as our ground truth for IRD patients, which is a laborious and time-consuming process. OCT scans were graded by two expert ophthalmologists. For IRDs, the segmentation of the outer retinal layers, especially the ONL, is essential to monitor structural degeneration over time. Thus, in our study, the focus was laid on a precise segmentation of the ONL. Within this work we will refer to the convex thickening of the ONL within the foveal pit as ONL-"hill" on central OCT slides, as depicted in ▶ **Fig. 3**.

Development of the deep learning-based segmentation algorithm and statistical analysis

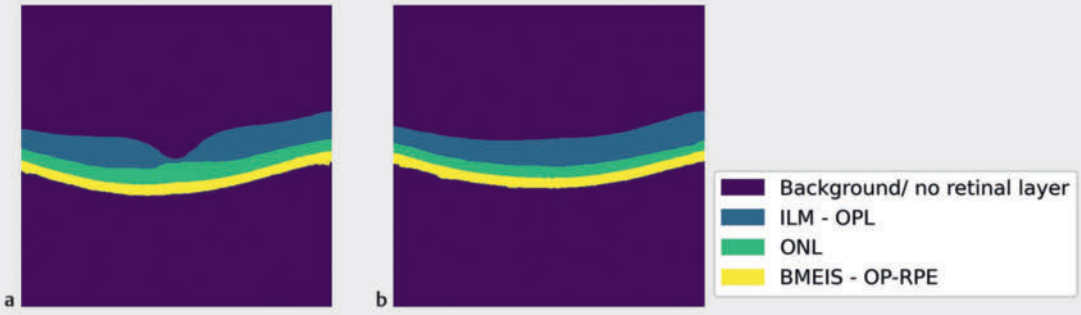
To the best of our knowledge, all existing deep learning-based approaches to segment retinal layers of OCT scans are based on a U-net architecture, which was also the choice for our model. Here, we rely on the Python module Tensorflow for our model definition and training. The versions of all packages we utilized can be found on GitHub in the requirements file. We ran the code within the de.NBI-cloud (Deutsches Netzwerk für Bioinformatik-Infrastruktur) to which we uploaded anonymized input images. The "de.NBI



▶ **Fig. 1** OCTExplorer segmentation screenshot for a healthy patient. Bold: segmented outer plexiform layer (OPL)-Henle fiber layer. OCT = optical coherence tomography; IRD = inherited retinal disease.



▶ **Fig. 2** OCTExplorer segmentation screenshot for an IRD patient. Bold: segmented outer plexiform layer (OPL)-Henle fiber layer. Note that the bold pink line for the segmentation of the OPL-Henle fiber layer boundary crosses over to the actual inner nuclear layer (white arrows). OCT = optical coherence tomography; IRD = inherited retinal disease.



▶ **Fig. 3** **a** Exemplary slide in the middle of an OCT scan in a healthy patient with the ONL-"hill". b Exemplary first slide of an OCT scan without ONL-"hill". The same color codes for retinal layers will be used within this work. OCT = optical coherence tomography; ONL = outer nuclear layer.

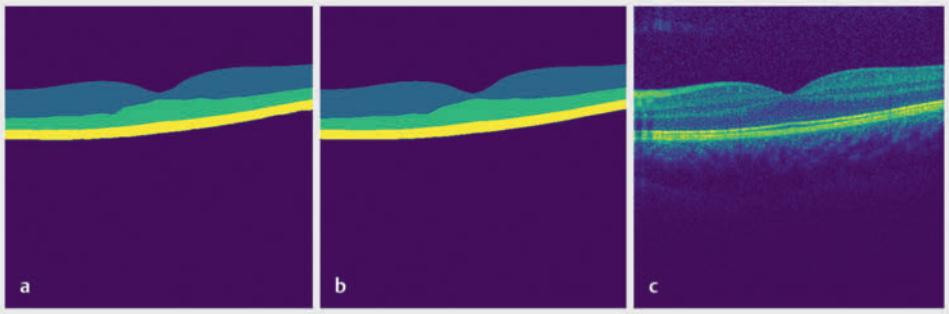
Cloud Berlin – Production" has provided us with sufficient compute instances for our project.

Since our OCT scan has an anisotropic resolution, we used 2D convolutional kernels rather than 3D kernels and processed the scan in a slide-by-slide fashion. Specifically, the backbone network consists of five symmetrical pairs of encoding and decoding blocks, which perform contraction and expansion operations, respectively, during the data feedforward process. Each encoding/decoding block is composed of two convolutional layers with (3 × 3) convolution kernels, followed by a batch normalization layer, an exponential linear unit (ELU), and a (2 × 2) max pooling operation with stride 2. The first encoding block results in 16 feature channels, and every subsequent encoding block doubles the number of channels. Each decoding block is symmetric with respect to the encoding block at the corresponding level, except that the max pooling layer is replaced with transposed convolutional layers to upsample the feature map. In the last decoding block, we added a (1 × 1) convolutional layer followed by the Soft-

▶ **Table 1** Pixel proportion of retinal layers of healthy and IRD patients datasets.

Retinal layer	Healthy OCTs	IRD OCTs
Background	83.8%	87.3%
ILM-OPL-HFL	8.8%	10.3%
ONL	4%	0.5%
BMEIS-OB-RPE	3.3%	1.8%

LM = internal limiting membrane,; OPL = outer plexiform layer; HFL = henle fibre layer; ONL = outer nuclear layer; BMEIS = Boundary of myoid and ellipsoid of inner segments; OB-RPE = outer boundary of retinal pigment epithelium



▶ **Fig. 4** Slide containing the ONL-"hill". **a** Prediction of the model. **b** Ground truth. **c** Input slide. ONL = outer nuclear layer.

max activation function to map the 16 feature channels to a four-class probability map that softly assigns each pixel to the background or one of the three foreground retinal layer classes, namely, the layers given in ▶ **Table 1**. The model architecture is very similar to [9], which has been shown to work for the task of brain tissue segmentation of different modalities/species. Instead of using a bottleneck dimension of 512, we have used a dimension of 1024. Further, we did not include the nonlocal attention block in our model, as this additional layer did not increase the model's performance in our task. Lastly, as data augmentation during training, we added random Gaussian noise to the input 2D slides.

Our objective loss function is a combination of the cross-entropy loss and the multiclass dice coefficient loss. For IRD patients, the ONL is much thinner compared to healthy subjects. Therefore, we introduced a different weight for different layers within the loss function, penalizing the ONL4 times higher than other layers for IRD patients, as our primary goal was to accurately segment the ONL. Here, we used one hot encoding of shape (4, 496, 512). The very last layer uses the Softmax activation function, returning the probability for each pixel belonging to one of the four classes.

To compensate for the limited manual annotations available for IRD patients, we utilized the wealth of automatic annotations from normal OCT scans generated by "OCTExplorer." By employing domain transfer and transfer learning techniques, we bridged the gap between IRD and standard OCT scans. This domain transfer network was originally developed by Yu et al. [9] for brain MRI segmentation, as described in detail in [9]. Essentially, we repurposed the weights of the model trained on healthy patients – we froze the weights of all trainable layers, except for all batch normalization (Batchnorm) layers, resulting in adaptive Batchnorm modules. By adapting this approach, we successfully achieved efficient learning using a small number of annotations and overcame the domain shifts between IRD and normal OCT scans. However, we also compared models that did not freeze the weights of all layers.

Results

Dataset characteristics

OCT scans from 12 healthy control individuals without retinal disease were collected, two OCT scans each, one for the left and right eye, respectively. For the healthy control training dataset, 18 volume scans consisting of 49 B-scans each from 12 patients met the quality criteria and were included. For the IRD training dataset, 16 patients were selected from the database. Because some OCT volume scans failed to meet the quality criteria, we included 25 volume scans in total, each consisting of either 49 or 97 B-scans.

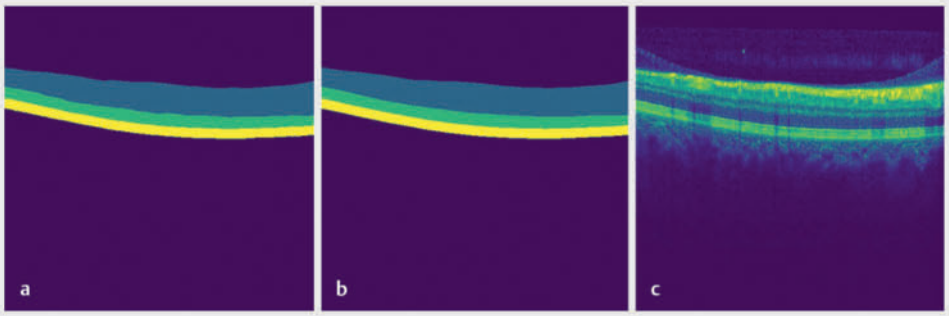
Due to the degenerative nature of IRDs, tissue loss also becomes evident by the unbalanced pixel proportion of the total scan area for different retinal layers, as shown in ▶ **Table 1**. Particularly, the ONL is considerably reduced compared to healthy controls (from 4.0% to only 0.5%).

For all results, we trained the U-net-based model for 30 epochs with early stopping. For all models trained on healthy or diseased patients, we used the same hold-out validation and test sets. Furthermore, we split training, validation, and test sets by patient IDs, such that there was no overlap between training and validation sets. Lastly, we used 2D slides as input to the model. The test and validation set consisted for both healthy and IRD datasets, each of scans of two patients' volume scans, mostly including a scan of the left and right eye, respectively. The codes and our trained models can be found on GitHub (https://github.com/peng-lab/retina-segmentation).

Results on healthy patients

First, we trained a model on the full training dataset, consisting of OCT scans of healthy individuals. Here, our model resulted in a dice coefficient of 99.4% for a hold-out test set. Exemplary segmentation results are illustrated in ▶ **Fig. 4** and **Fig. 5**. It is noteworthy that we just included OCT scans that were labeled and verified by our expert annotators.

We also trained models containing the nonlocal block, as described in [9], but it does not outperform the model not including this additional layer quantitatively. Further, it led to occasionally isolated misclassified pixels. Eventually, we decided to not include the nonlocal block.



▶ **Fig. 5** Regular slide. **a** Prediction of the model. **b** Ground truth. **c** Input slide.

▶ **Table 2** Dice coefficients on the IRD dataset.

Model	Description	Dice coefficient	Dice coefficient ONL
Model 1	Non-weighted ONL loss	98.5%	44.2%
Model 2	Weighted ONL loss	98.5%	67.4%
Model 3	Direct application of healthy model	92.1%	31.7%
Model 4	DA by freezing trainable weights	98.8%	70.2%
Model 5	Transfer learning without freezing weights	98.8%	70.1%

DA = domain adaption; ONL = outer nuclear layer

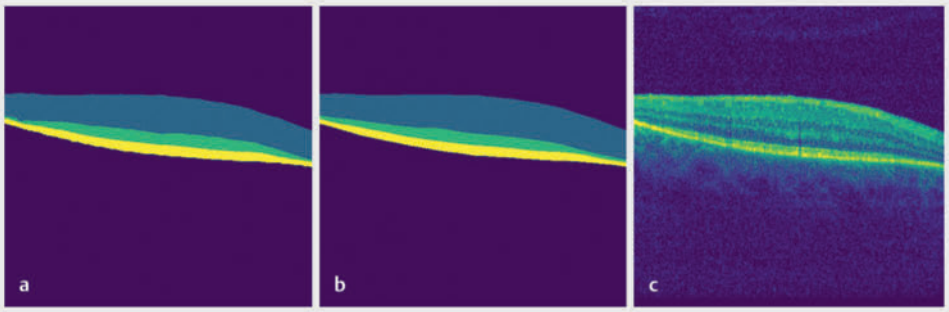
Results on IRD patients

To segment retinal layers of IRD patients, we tried a few different approaches: Model 1 trained with the same loss function as for healthy patients with randomly initialized weights, (2) Model 2 trained by upweighting the ONL in the cross-entropy loss with randomly initialized weights, (3) Model 3 directly applying the learned model of healthy subjects on IRD patients, (4) Model 4 using domain adaption (DA) by freezing all weights apart from the batch normalization layers of the model trained on healthy subjects (adaptive Batchnorm), and (5) Model 5 trained with transfer learning by reusing the weights of the model trained on healthy individuals for initialization, however, without freezing the weights of all layers. Notably, models reusing weights from the previous model were fitted faster. The corresponding dice coefficients can be found in ▶ **Table 2**.

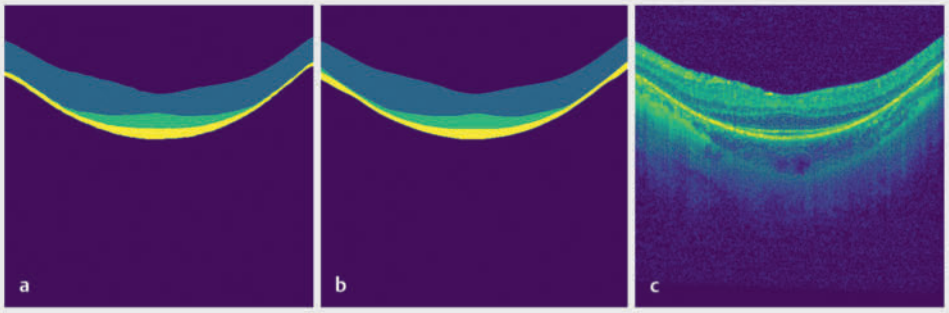
Model 1 failed to correctly segment the middle slides, with the convex curvature of the ONL within the foveal pit. Therefore, we introduced a weighted loss for the ONL for all remaining models. We found that a weight of 4 provides the best results to accurately segment the ONL. This determination was made through hyperparameter optimization. From Model 2, we concluded that upweighting the ONL in the cross-entropy loss helped to better capture and segment this layer and it further increased the overall dice coefficient. Without using the weighted loss, the model oftentimes did not capture the ONL at all for IRD patients, as this layer is nearly nonexistent in many slides. Therefore, the model ended up not assigning any probability to the ONL.

Furthermore, we could conclude that we needed two distinct models to segment OCT scans of IRD and healthy individuals. Directly applying the model trained on healthy individuals, i.e., Model 3, lead to poor performance compared to other models. The models that exhibited the highest performance were those that upweighted the ONL within the cross-entropy loss. However, by using the already learned features from the model trained on healthy control individuals, the model could increase its performance. Models 4 and 5 exhibited the highest dice scores while they were simultaneously trained more quickly in comparison to Models 1 and 2. The optimal performance was attained by the DA model, which uses the adaptive Batchnorm approach. Consequently, we opted to use this model for visualization purposes. ▶ **Fig. 6**, **7**, and **8** show segmentation results for input slides from our test set. Notably, if the quality of the scan was low, the model occasionally ended up misclassifying a few random pixels, as depicted in ▶ **Fig. 9**.

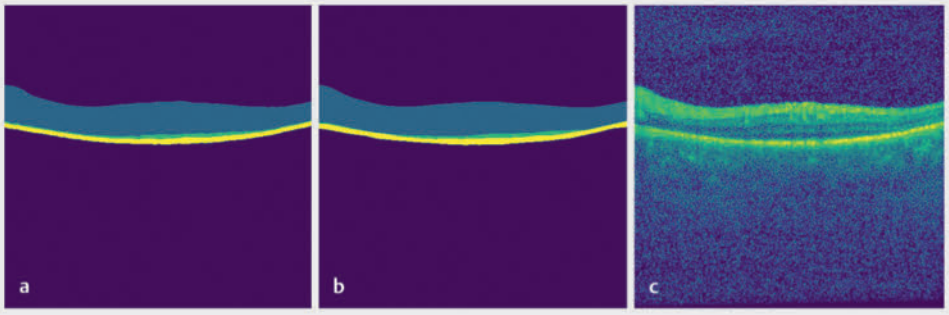
We also worked on a grid-based visual analysis using the widely used ETDRS (Early Treatment Diabetic Retinopathy Study) regions as depicted in ▶ **Fig. 10**. Based on the segmentation results, we calculated thickness maps of all retinal layers or of individual retinal layers. Exemplary thickness maps including the ETDRS regions/grid are depicted in ▶ **Fig. 11** and **12**.



▶ **Fig. 6** Trained model on the IRD dataset. **a** Prediction of the model. **b** Ground truth. **c** Input slide. IRD = inherited retinal disease.



▶ **Fig. 7** Trained model on the IRD dataset where ONL-"hill" is visible. **a** Prediction of the model. **b** Ground truth. **c** Input slide. IRD = inherited retinal disease; ONL = outer nuclear layer.

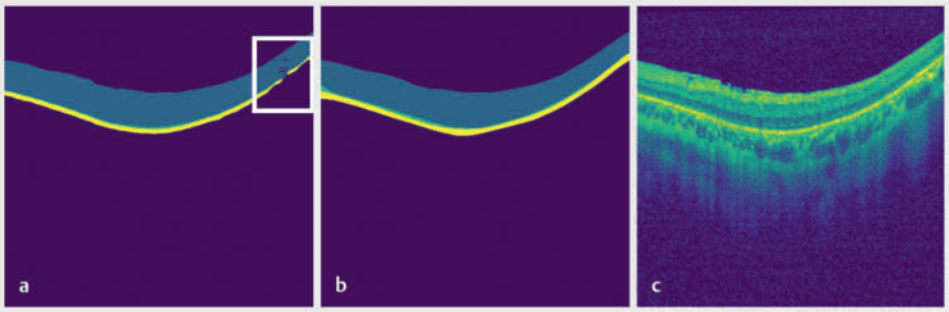


▶ **Fig. 8** Trained model on the IRD dataset where the ONL is almost vanished. **a** Prediction of the model. **b** Ground truth. **c** Input slide. IRD = inherited retinal disease; ONL = outer nuclear layer.

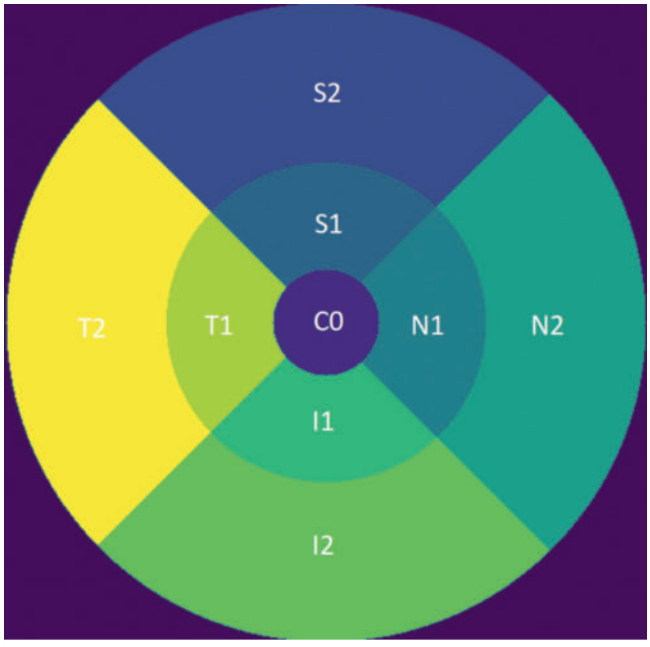
Discussion

The aim of the present study was to build an AI-based tool for the segmentation of OCTs in IRD patients, where currently available algorithms fail. Additionally, we developed a full retina and ONL thickness map, which provides a fast overview of the important retinal layers in IRD.

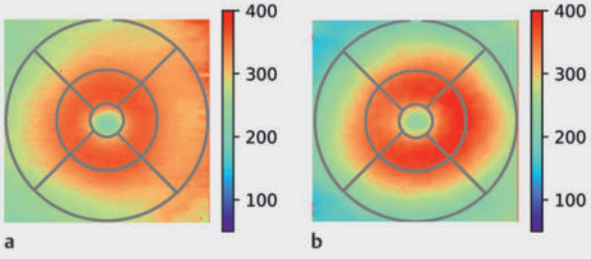
With the advent of AI and machine learning, several studies aiming to enhance diagnosis, prognosis, and treatment options for ophthalmological diseases have been published [10 – 14]. Evaluation of retinal imaging is essential in the diagnosis of these diseases. It is a time-consuming process that requires specialists and may show variable interpretation depending on the examiner [14]. OCT allows for noninvasive structural retinal imaging and



▶ **Fig. 9** Trained model on an IRD dataset. The model has problems when the quality is too low. White frame indicates some wrongly segmented pixels. **a** Prediction of the model. **b** Ground truth. **c** Input slide. IRD = inherited retinal disease.



▶ **Fig. 10** ETDRS regions.



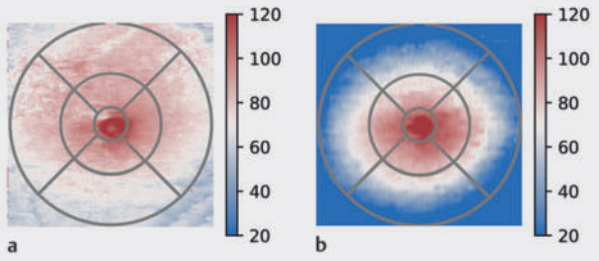
▶ **Fig. 11** Thickness maps based on predicted segmentation of all retinal layers. **a** Thickness map of a healthy individual. **b** Thickness map of an IRD patient. Interestingly, the diseased individual has a similar full retinal thickness. Healthy: C0 average thickness: 274 µm; S2 average thickness: 301 µm; S1 average thickness: 354 µm; N1 average thickness: 352 µm; N2 average thickness: 325 µm; I1 average thickness: 352 µm; I2 average thickness: 302 µm; T1 average thickness: 348 µm; T2 average thickness: 294 µm. Diseased: C0 average thickness: 281 µm; S2 average thickness: 270 µm; S1 average thickness: 359 µm; N1 average thickness: 369 µm; N2 average thickness: 310 µm; I1 average thickness: 366 µm; I2 average thickness: 279 µm; T1 average thickness: 343 µm; T2 average thickness: 278. IRD = inherited retinal disease.

generates important information about disease development and treatment response. AI-based accurate OCT analysis could be a huge advantage and potentially prevent individual interpretation errors, depending on the investigator.

However, compared to more common ophthalmological diseases, AI studies specialized on IRDs are still scarce. A systemic PubMed search found only two existing works for deep learning-based OCT studies on IRDs. Camino et al. [15] investigated the deep learning-based segmentation of preserved photoreceptors on OCT images in the two IRD subtypes choroideremia and retinitis pigmentosa. Zhao et al. [16] developed few-shot learning approaches to classify OCT images of IRDs into disease subtypes. Miere et al. [17] used a convolutional neural network (CNN) to classify IRDs into subtypes by analyzing fundus autofluorescence images.

Deep learning algorithms usually require a huge training set, or at least a sufficient number of train images, to result in good generalizable segmentations. However, in our study, we exclusively incorporated meticulously chosen OCT volumes, which were characterized by their high quality in terms of having precise annotations. Therefore, we reason that our model performs very well on unseen data, as it was trained on two high-quality datasets (healthy and IRD). This achievement was made possible due to the close collaboration between clinicians and AI experts. Subsequently, we could rely on a model [9] that has been shown to work well without resorting to more complex architectures or deeper networks, clearly highlighting the strength of our results.

OCT imaging has led to an abrupt and intense advancement of ophthalmological research, as it enabled retinal *in vivo* imaging in a resolution that was formerly only known from histological tissue



▶ **Fig. 12** Thickness maps based on predicted segmentation of the ONL. **a** Thickness map of a healthy individual. **b** Thickness map of an IRD patient. Evidently, the IRD patient has a thinner ONL thickness, especially at the non-central regions. Healthy: C0 average thickness: 109 µm; S2 average thickness: 80 µm; S1 average thickness: 88 µm; N1 average thickness: 94 µm; N2 average thickness: 75 µm; I1 average thickness: 88 µm; I2 average thickness: 68 µm; T1 average thickness: 89 µm; T2 average thickness: 76 µm. Diseased: C0 average thickness: 115 µm; S2 average thickness: 37 µm; S1 average thickness: 81 µm; N1 average thickness: 94 µm; N2 average thickness: 48 µm; I1 average thickness: 92 µm; I2 average thickness: 42 µm; T1 average thickness: 89 µm; T2 average thickness: 51 µm. IRD = inherited retinal disease.

microscopy. While structural abnormalities such as intraretinal fluid accumulations are monitored in exudative retinal diseases, in IRDs, the focus lies on monitoring slowly progressive tissue loss, which can be followed up over years. In IRDs of the retinitis pigmentosa type, the outer retinal layers are affected, while the inner retinal layers are typically preserved. For IRDs, the segmentation of the outer retinal layers is essential to monitor structural degeneration over time and has also been applied in animal models of preclinical studies [18]. The preservation of outer retinal layers, especially the ONL, in OCT scans has been shown to correlate with residual vision in dark-adapted visual fields in AMD patients [19]. In a study of patients with *RPE65*-related IRD by Jacobson and colleagues, 96% of retinal loci with residual light sensitivity had a measurable ONL while in 75% of retinal loci without residual sensitivity, no ONL was measurable [20]. Not only the "OCTExplorer" used for this study, but also most other commercially available automated segmentation tools, fail to reliably segment the distinct layers in OCT scans of patients with IRDs, which results in time-consuming manual correction and/or segmentation [19, 21].

In 2017, voretigene neparvovec-rzyl (Luxturna), an adeno-associated virus (AAV) vector for gene augmentation therapy in patients with *RPE65*-related IRD, reached its primary endpoint in a phase III trial, which marked a breakthrough in the field of gene therapy [22]. However, since then, other phase III trials on ocular gene therapy have failed to reach their primary endpoints, e.g., the XIRIUS study of an AAV serotype 8 (AAV8) vector-based gene therapy, cotoretigene toliparvovec, targeting the RPGR gene in X-linked retinitis pigmentosa. One of the recurring key questions in all ophthalmological gene therapy studies therefore is: "How to measure success?" [23]. This question underlines the complexity and importance of a reliable and objective measurement of visual performance for gene therapy studies in IRDs and beyond. A huge advantage in the diagnostic toolbox for IRDs would be a correlation of structural OCT measurements and functional data. For structure-function prediction, reliable retinal layer segmentation is regarded as crucial. Hence, segmentation algorithms like the one presented in this manuscript could play an important role in future studies.

There are limitations to this study. In contrast to AMD or DR, IRDs are rare diseases, which limits the sample size. Training of the AI was based on manual segmentation of image stacks by experienced ophthalmologists. Manual segmentation bears the risk of human error and may differ slightly between graders. To manually segment the OCT scans for AI training datasets, a minimum resolution quality is needed. However, in real-life settings, imaging resolution in IRD patients may be low due to fixation problems and blinking artifacts. The present study was only performed with Heidelberg Spectralis OCT (Heidelberg Engineering, Version 1.10.12.0) scans, further investigations should confirm this AI-based tool with scans of other commercially available devices.

In conclusion, AI-based tools are thought to be the key to boost efficiency for an objective assessment of rare diseases, where well-powered randomized controlled trials are almost impossible to conduct due to the limited sample size. The here presented segmentation and thickness map tool may be a first step to AI-supported IRD diagnostics. It could enable structural disease monitoring in future gene therapy trials and provide the basis for predictive structure-functional modelling.

CONCLUSION BOX

Already known:

- AI has an increasingly important role in ophthalmology. To date, most research has focused on the high prevalence ophthalmic diseases.
- Our study addresses a critical relevant problem, focusing on reliable retinal layer segmentation for IRD patients. Accurate segmentation of anatomical layers in OCT scans plays a key role in the correlation of retinal structure to visual function.

Newly described:

- We have developed a deep learning algorithm that allows accurate segmentation of pathologically altered OCT scans in patients with IRDs and generates a retinal thickness map.
- Future work will explore calculating retinal layer thickness and correlating it with functional data, like visual fields. Our aim is to contribute to a greater understanding of disease, and to improve the evaluation of treatment outcomes in the future.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Aleman TS, Cideciyan AV, Sumaroka A et al. Retinal laminar architecture in human retinitis pigmentosa caused by *Rhodopsin* gene mutations. Invest Ophthalmol Vis Sci 2008; 49: 1580–1591. doi:10.1167/IOVS.07-1110

[2] Apushkin MA, Fishman GA, Alexander KR et al. Retinal thickness and visual thresholds measured in patients with retinitis pigmentosa. Retina 2007; 27: 349–357. doi:10.1097/01.IAE.0000224944.33863.18

[3] Abràmoff MD, Lavin PT, Birch M et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018; 1: 39. doi:10.1038/s41746-018-0040-6

[4] Lee CS, Baughman DM, Lee AY. Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images. Ophthalmol Retin 2017; 1: 322–327. doi:10.1016/J.ORET.2016.12.009

[5] Li Z, He Y, Keel S et al. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. Ophthalmology 2018; 125: 1199–1206. doi:10.1016/J.OPHTHA.2018.01.023

[6] Abramoff MD, Garvin MK, Sonka M. Retinal imaging and image analysis. IEEE Rev Biomed Eng 2010; 3: 169–208. doi:10.1109/RBME.2010.2084567

[7] Li K, Wu X, Chen DZ et al. Optimal surface segmentation in volumetric images–a graph-theoretic approach. IEEE Trans Pattern Anal Mach Intell 2006; 28: 119. doi:10.1109/TPAMI.2006.19

[8] Garvin MK, Abràmoff MD, Wu X et al. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. IEEE Trans Med Imaging 2009; 28: 1436–1447. doi:10.1109/TMI.2009.2016958

[9] Yu Z, Han X, Xu W et al. A generalizable brain extraction net (BEN) for multimodal MRI data from rodents, nonhuman primates, and humans. Elife 2022; 11: e81217. doi:10.7554/ELIFE.81217

[10] Ting DSW, Pasquale LR, Peng L et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol 2019; 103: 167–175. doi:10.1136/BJOPHTHALMOL-2018-313173

[11] Moraes G, Fu DJ, Wilson M et al. Quantitative Analysis of OCT for Neovascular Age-Related Macular Degeneration Using Deep Learning. Ophthalmology 2021; 128: 693–705. doi:10.1016/J.OPHTHA.2020.09.025

[12] Schmidt-Erfurth U, Vogl WD, Jampol LM et al. Application of Automated Quantification of Fluid Volumes to Anti–VEGF Therapy of Neovascular Age-Related Macular Degeneration. Ophthalmology 2020; 127: 1211–1219. doi:10.1016/j.ophtha.2020.03.010

[13] Hogarty DT, Mackey DA, Hewitt AW. Current state and future prospects of artificial intelligence in ophthalmology: a review. Clin Experiment Ophthalmol 2019; 47: 128–139. doi:10.1111/CEO.13381

[14] Balyen L, Peto T. Promising artificial intelligence–machine learning–deep learning algorithms in ophthalmology. Asia Pac J Ophthalmol (Phila) 2019; 8: 264–272. doi:10.22608/APO.2018479

[15] Camino A, Wang Z, Wang J et al. Deep learning for the segmentation of preserved photoreceptors on en face optical coherence tomography in two inherited retinal diseases. Biomed Opt Express 2018; 9: 3092. doi:10.1364/BOE.9.003092

[16] Zhao Q, Mai SW, Li Q et al. Automated Classification of Inherited Retinal Diseases in Optical Coherence Tomography Images Using Few-shot Learning. Biomed Environ Sci 2023; 36: 431–440. doi:10.3967/BES2023.052

[17] Miere A, Le Meur T, Bitton K et al. Deep Learning-Based Classification of Inherited Retinal Diseases Using Fundus Autofluorescence. J Clin Med 2020; 9: 1–13. doi:10.3390/JCM9103303

[18] Gardiner KL, Cideciyan AV, Swider M et al. Long-Term Structural Outcomes of Late-Stage *RPE65* Gene Therapy. Mol Ther 2020; 28: 266–278. doi:10.1016/J.YMTHE.2019.08.013

[19] Kelbsch C, Stingl K, Kempf M et al. Objective Measurement of Local Rod and Cone Function Using Gaze-Controlled Chromatic Pupil Campimetry in Healthy Subjects. Transl Vis Sci Technol 2019; 8: 19. doi:10.1167/TVST.8.6.19

[20] Jacobson SG, Aleman TS, Cideciyan AV et al. Defining the residual vision in Leber congenital amaurosis caused by RPE65 mutations. Invest Ophthalmol Vis Sci 2009; 50: 2368–2375. doi:10.1167/IOVS.08-2696

[21] Gersch J, Hufendiek K, Delarocque J et al. Investigation of Structural Alterations in Inherited Retinal Diseases: A Quantitative SD-OCT-Analysis of Retinal Layer Thicknesses in Light of Underlying Genetic Mutations. Int J Mol Sci 2022; 23: 16007. doi:10.3390/IJMS232416007/S1

[22] Russell S, Bennett J, Wellman JA et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. Lancet 2017; 390: 849. doi:10.1016/S0140-6736(17)31868-8

[23] Stingl K, Kempf M, Jung R et al. Therapy with voretigene neparvovec. How to measure success? Prog Retin Eye Res 2023; 92: 101115. doi:10.1016/J.PRETEYERES.2022.101115