AMERICAN ACADEMY OF OPHTHALMOLOGY®

Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning

Thomas Schlegl, MSc,[1,4] Sebastian M. Waldstein, MD, PhD,[2] Hrvoje Bogunovic, PhD,[1] Franz Endstraßer, MD,[1] Amir Sadeghipour, PhD,[1] Ana-Maria Philip, MD,[3] Dominika Podkowinski, MD,[3] Bianca S. Gerendas, MD, MSc,[2] Georg Langs, PhD,[4] Ursula Schmidt-Erfurth, MD[1,2]

Purpose: Development and validation of a fully automated method to detect and quantify macular fluid in conventional OCT images.

Design: Development of a diagnostic modality.

Participants: The clinical dataset for fluid detection consisted of 1200 OCT volumes of patients with neovascular age-related macular degeneration (AMD, n = 400), diabetic macular edema (DME, n = 400), or retinal vein occlusion (RVO, n = 400) acquired with Zeiss Cirrus (Carl Zeiss Meditec, Dublin, CA) (n = 600) or Heidelberg Spectralis (Heidelberg Engineering, Heidelberg, Germany) (n = 600) OCT devices.

Methods: A method based on deep learning to automatically detect and quantify intraretinal cystoid fluid (IRC) and subretinal fluid (SRF) was developed. The performance of the algorithm in accurately identifying fluid localization and extent was evaluated against a manual consensus reading of 2 masked reading center graders.

Main Outcome Measures: Performance of a fully automated method to accurately detect, differentiate, and quantify intraretinal and SRF using area under the receiver operating characteristics curves, precision, and recall.

Results: The newly designed, fully automated diagnostic method based on deep learning achieved optimal accuracy for the detection and quantification of IRC for all 3 macular pathologies with a mean accuracy (AUC) of 0.94 (range, 0.91−0.97), a mean precision of 0.91, and a mean recall of 0.84. The detection and measurement of SRF were also highly accurate with an AUC of 0.92 (range, 0.86−0.98), a mean precision of 0.61, and a mean recall of 0.81, with superior performance in neovascular AMD and RVO compared with DME, which was represented rarely in the population studied. High linear correlation was confirmed between automated and manual fluid localization and quantification, yielding an average Pearson's correlation coefficient of 0.90 for IRC and of 0.96 for SRF.

Conclusions: Deep learning in retinal image analysis achieves excellent accuracy for the differential detection of retinal fluid types across the most prevalent exudative macular diseases and OCT devices. Furthermore, quantification of fluid achieves a high level of concordance with manual expert assessment. Fully automated analysis of retinal OCT images from clinical routine provides a promising horizon in improving accuracy and reliability of retinal diagnosis for research and clinical practice in ophthalmology. *Ophthalmology 2018;125:549-558 © 2017 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.aaojournal.org.

OCT has profoundly disrupted conventional diagnostic and therapeutic strategies in clinical management and has led to paradigm shifts in the understanding of macular disease. Although OCT has continuously undergone hardware improvements since its inception,[1] significantly less progress has been made in the field of software methods to analyze clinical OCT data. The number of patients with macular disease requiring efficient disease management based on OCT in clinical practice continues to increase, similarly to the amount of image data produced by advanced OCT technology such as swept source. Therefore, the feasibility of manual OCT assessment in clinical practice has become largely unrealistic. Likewise, poor reproducibility between OCT assessors, even in a research setting, also has been reported.[2] Consequently, automated and reproducible analysis of clinical OCT data represents an important unmet need and a promising perspective for clinical practice. Specifically, there is a clear need to advance automated analysis beyond a purely anatomic presence/absence detection to an accurate measurement of markers for disease activity.

The majority of available analysis software tools for OCT is limited to the measurement of retinal (layer) thickness, despite the fact that prior studies demonstrated the limited value of this biomarker for visual prognosis and disease management.[3] In practice, most physicians use qualitative OCT biomarkers, such as the presence of intraretinal cystoid fluid (IRC) and subretinal fluid (SRF), to inform retreatment decisions in individualized

https://doi.org/10.1016/j.ophtha.2017.10.031
ISSN 0161-6420/17

therapy of macular diseases. These biomarkers are also sophisticated prognostic factors for visual outcome, with IRC representing 1 of the most important variables associated with vision loss and SRF possibly enhancing visual prognosis.[4]

Prior studies have proposed to detect or quantify macular fluid in OCT in an automated manner.[5] However, they are limited to only being able to detect fluid presence or absence without measuring its extent and distribution, by a lack of differentiation between IRC and SRF, or by evaluation limited to 1 particular OCT device or disease.

In this study, we present a fully automated artificial intelligence method to detect and quantify IRC and SRF in macular OCT volume scans. We validate our method on a large dataset of eyes presenting with the major relevant exudative macular diseases, that is, neovascular age-related macular degeneration (AMD), diabetic macular edema (DME), and macular edema secondary to retinal vein occlusion (RVO), imaged with the most commonly used OCT devices (Heidelberg Spectralis, Heidelberg Engineering, Heidelberg, Germany, and Zeiss Cirrus, Carl Zeiss Meditec, Dublin, CA).

Methods

Image Dataset

This study followed the tenets set forth in the Declaration of Helsinki, and approval was obtained by the Ethics Committee of the Medical University of Vienna. A total of 1200 completely anonymized OCT volume scans of eyes affected by the major diseases typically causing macular fluid (AMD, DME, and RVO) were extracted from the Vienna Reading Center database. Furthermore, for each disease we selected scans by 2 different OCT devices (Cirrus HD-OCT, Carl Zeiss Meditec, and Spectralis OCT, Heidelberg Engineering), resulting in 6 distinct groups (disease × device) of OCT scans.

For detection of fluid, we used 1200 anonymized OCT volumes of patients with neovascular AMD (n = 400), DME (n = 400), and RVO (n = 400) acquired with Cirrus (50%) or Spectralis (50%) devices. In each disease/device group, 50% of the OCT scans showed retinal fluid.

All scans were graded for the presence of IRC and SRF in a quality-controlled reading center setting by 2 independent readers of the Vienna Reading Center supervised by experienced retinal specialists. The grading was performed on full-screen, high-resolution 27-inch monitors in validated grading software specially designed for the annotation of these features. In questionable cases, a consensus grading of the readers and the supervisor was performed. Only scans with a clear consensus annotation between the readers were taken into the sample. Scans with primarily low image quality (i.e., scans with sections cut off because of improper positioning during image acquisition or scans with strong motion artifacts causing misalignment and blurring of sections) were excluded. We randomly selected 100 scans with fluid and 100 scans without fluid per disease/device group from the OCT database, resulting in a total of 1200 scans.

For evaluation of fluid quantification, datasets with complete manual annotations of IRC and SRF that were available at the Vienna Reading Center were randomly selected. For neovascular AMD, we used a Cirrus dataset of 152 scans as reported previously[6] and a Spectralis dataset consisting of 60 scans. For DME, we used a dataset of 16 Cirrus scans and a dataset of 16 Spectralis scans. For RVO, we used a Cirrus dataset of 100 scans and a Spectralis dataset consisting of 10 scans. The procedure used for manual annotation of macular fluid has been reported in detail by Waldstein et al.[6] Each Cirrus cube scan consisted of 128 B-scans with a resolution of 512 × 1024 pixels, with the exception of Cirrus RVO data, which consisted of 200 B-scans with a resolution of 200 × 1024 pixels. All cube scans acquired on the Spectralis device consisted of 49 B-scans at a resolution of 512 × 496 pixels. Both machines acquire volumes covering approximately 6 × 6 × 2 mm^3 corresponding to a field of view of 20° × 20°.

Description of the Automated Method

We developed a software to perform classification among IRC, SRF, and nonfluid regions for each location (= pixel) in the OCT image by using and further developing semantic segmentation,[7] a method based on convolutional neural networks. Specifically, we applied deep learning, a state-of-the-art machine learning technique in the field of artificial intelligence that learns the mapping from OCT images to pixel-level class labels based on large amounts of labeled training data. Deep learning models allow one to learn meaningful abstract data representations. Following the semantic segmentation approach, the neural network maps an input image of a specific size to an image of corresponding class labels of the same size. The proposed neural network comprises 2 processing components, an encoder that transforms an input image into an abstract representation and a decoder that maps the abstract representation to an image of clinical class labels assigning each pixel a class such as normal tissue, IRF, or SRF.

The mapping of the encoder from raw images to abstract representations (embeddings) was not computed on the basis of prespecified mathematic descriptions (handcrafted features), but the encoder parameters were automatically learned solely on the basis of annotated data used during training. The data embedding learned was optimized in such a way that it was optimal for the generation of a corresponding image of class labels. The mapping of the encoder from raw images to the data embedding needed to generate the label image, and the mapping of the decoder from the embedding to a full input resolution label image were learned simultaneously (end-to-end). The encoder and the decoder comprised a set of computing blocks (layers), where the layers of the decoder virtually inversed the operations of the *encoder* conditioned by the low-dimensional embedding learned by the encoder. A simplified overview of the encoder-decoder architecture is shown in Figure 1, and a detailed description of the algorithm appears in the Appendix (available online at www.aaojournal.org).

The runtime of the computation depended on the resolution of OCT volumes. For Cirrus scans, the computations took approximately 70 seconds, whereas for Spectralis scans, computations took approximately 30 seconds per volume on a TitanX graphical processing unit (Nvidia, Santa Clara, CA) with NVIDIA Cuda (version 6.5) and cuDNN library (version 2).[8] Python, Theano,[9] and the Lasagne library were used to train and evaluate the deep learning model.[10]

Experimental Setup

Because of differences in the appearance of Spectralis and Cirrus scans, we trained separate classifiers for Spectralis and Cirrus scans. We used 70 pixel-wise annotated OCT scans performing 10-fold cross-validation including 7 scans per split for training and testing of the model on Spectralis data. We performed 4-fold cross-validation for training and testing of the model on Cirrus data, where we had 257 pixel-wise annotated OCT scans and used 64 and 65 scans per split. Both datasets comprised neovascular AMD and RVO cases. When considerably fewer scans were used to train the model, results showed that the proposed method also achieved a similar performance. However, the achievable performance, and at the same time the applicability for clinical routine, increased with the number of training volumes. When the model trained on Cirrus data was used on Spectralis data and vice versa, results on quantifying IRC or SRF showed that the model trained on Cirrus

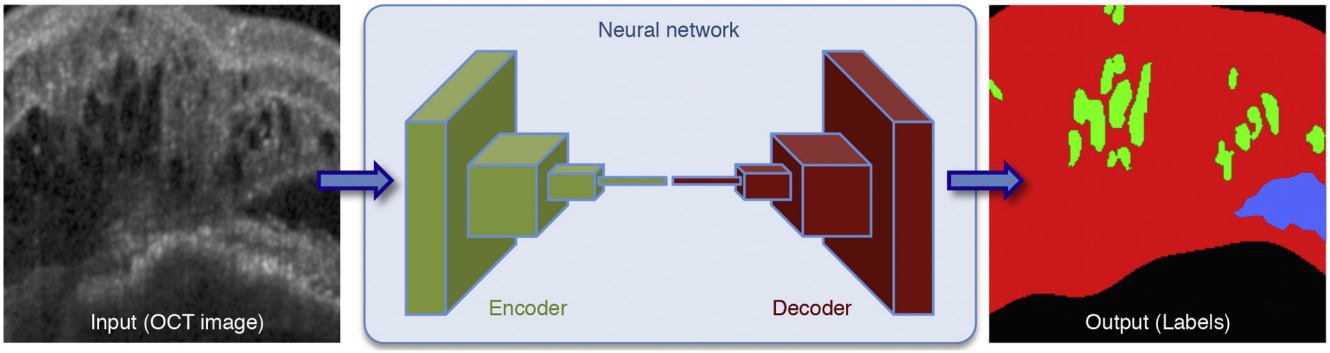

Figure 1. Illustration of the proposed automated method. The approach comprises a convolutional neural network with an *encoder-decoder* architecture to identify intraretinal cystoid fluid (IRC) and subretinal fluid (SRF). The *encoder* maps an OCT image to an abstract representation (*embedding*). The *decoder*, in turn, maps this embedding to a full input resolution label image. The IRC is marked in *green*, SRF is marked in *blue*, and nonfluid retinal tissue is marked in *red*. Background areas are marked in *black*.

data obtained a similar performance when applied on Spectralis images. However, a model trained on Spectralis data showed loss of performance when applied on Cirrus data.

Because of differences in class label distributions of DME cases compared with neovascular AMD or RVO cases, we did not include DME cases in the dataset for classifier training. Instead, we used 16 DME cases to adapt the model trained on neovascular AMD and RVO to the class label distribution of DME cases. We used separate DME scans for Spectralis and Cirrus models.

In each run of the cross-validation, we extracted a set of pixels, image patches that show the morphologic context of the pixel, and the corresponding annotated label (normal, SRF, IRF). This set was used for training of the neural network. On the test set, we applied the trained network and compared the resulting pixel-wise labeling with the expert annotator labels. We performed exhaustive pixel sampling, resulting in every pixel getting multiple neural-network class votes. The final label (normal, SRF, IRF) was determined by majority vote.

Because algorithms often fail with macular fluid, we used only the ILM and RPE layers, which are mostly robustly segmented.[11] The distinction of individual retinal layers, which are indeed erroneous when pronounced macular fluid is present, is not required. Thus, we only had to deal with few erroneous cases. These cases exhibited single positions (columns) within B-scans where the layer segmentation yielded undefined values (erroneously denoting layer positions on the outside of the image plane), which were easily and reliably corrected via interpolation between valid neighboring values.

The model was trained to predict the correct class label for every pixel position within the image patch. During testing, we exhaustively extracted partially overlapping image patches and used majority voting to obtain single class labels for every pixel of the retina. This resulted in a dense segmentation of the retinal region of an OCT scan so that every pixel showing the retina received 1 class label (IRC, SRF, or healthy retinal tissue).

Statistical Evaluation

We evaluated quantification (pixel-wise labeling of an image) and detection (determining if a volume contains SRF, IRF, or no fluid). The evaluation of both the volume-level detection performance and the pixel-level segmentation accuracy was performed on the basis of pixel-wise IRC and SRF segmentations by the software and corresponding ground truth annotations by reading experts. For volumes on which the model was trained and the quantification performance was evaluated, we had pixel-wise ground truth annotations. In contrast, only volume-level ground truth labels were available for volumes on which detection performance was evaluated.

Table 1. Detection of Fluid Presence

		Cirrus (Carl Zeiss Meditec, Dublin, CA)			Spectralis (Heidelberg Engineering, Heidelberg, Germany)		
Fluid Type	**Measure**	*AMD*	*DME*	*RVO*	*AMD*	*DME*	*RVO*
IRC	Precision	0.75	0.86	0.97	0.92	0.99	0.95
	Recall	0.92	0.90	0.77	0.71	0.89	0.86
	AUC	0.93	0.93	0.94	0.91	0.97	0.95
	n	64	99	100	65	100	100
SRF	Precision	0.93	0.26	0.73	0.95	0.67	0.66
	Recall	0.93	0.67	0.81	0.90	0.55	0.88
	AUC	0.98	0.90	0.92	0.98	0.87	0.98
	n	81	9	54	69	11	26

AMD = age-related macular degeneration; AUC = area under the curve; DME = diabetic macular edema; IRC = intraretinal cystoid fluid; RVO = retinal vein occlusion; SRF = subretinal fluid.
Precision, recall, and AUC of volume-level detection of IRC and SRF evaluated on neovascular AMD, DME, and RVO cases at the Youden Index of the corresponding receiver operating characteristic (ROC) curve. The "n" values indicate the number of volumes per disease/machine group containing IRC or SRF, respectively.

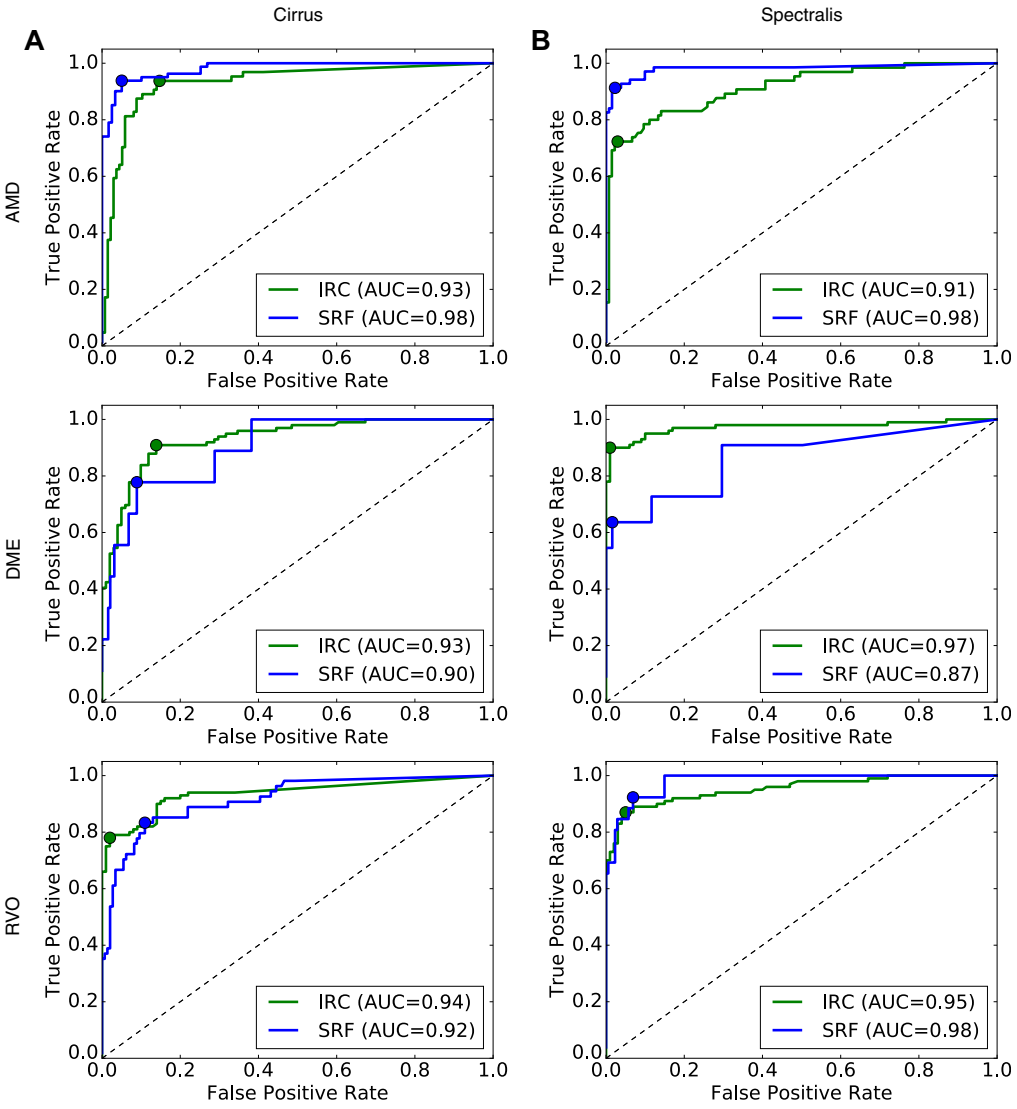

Figure 2. Receiver operating characteristic (ROC) curves on volume-wise detection performance of intraretinal cystoid fluid (IRC) and subretinal fluid (SRF). The Youden Index is indicated by the *dot*. **First row**: Neovascular age-related macular degeneration (AMD). **Second row**: Diabetic macular edema (DME). **Bottom row**: Retinal vein occlusion (RVO). **A**, Cirrus data (Carl Zeiss Meditec, Dublin, CA). **B**, Spectralis data (Heidelberg Engineering, Heidelberg, Germany). The area under the curve (AUC) as a measure of general performance is specified in parentheses.

Validation of the Detection of Retinal Fluid

For the binary classification of whether a volume comprises fluid, we used a threshold over the number of fluid pixels. To evaluate the overall detection performance, we computed receiver operating characteristic (ROC) curves by varying the threshold over the number of fluid pixel segmented by our model.

We evaluated the retinal fluid detection performance on neovascular AMD, DME, and RVO cases on Cirrus and Spectralis data. For each of these disease/device groups, ROC curves were computed separately for IRC and SRF. On the basis of the ROC curves, we computed the area under the curve (AUC). An AUC can be interpreted as the mean sensitivity value for all possible specificity values, or equivalently, as the mean specificity value for all possible sensitivity values (https://analyse-it.com/docs/user-guide/diagnosticperformance/auc). The possible AUC value ranges from 0.50 (discriminative performance equal to chance) to 1.00 (perfect discriminative performance).

Furthermore, we computed precision and recall of volume-based detection of IRC and SRF evaluated at the Youden Index,[12] an optimal cutoff point that simultaneously maximizes sensitivity and specificity of the corresponding ROC curve. Also, we report the number of volumes containing IRC or SRF.

Validation of Retinal Fluid Quantification

For clinical management, not only the presence of IRC or SRF within a volume scan but also the amount of fluid and its exact location are of interest. Thus, we evaluated the pixel-level segmentation performance of our proposed method. Again, the evaluation of retinal fluid segmentation accuracy was performed independently for IRC and SRF and evaluated on neovascular AMD, DME, and RVO cases on Cirrus and Spectralis data. The segmentation performance per disease/device group was computed as average over volume-wise precision and recall of pixel-wise segmentations of IRC and SRF. Volumes that did

Table 2. Quantification of Retinal Fluid

Fluid Type	Measure	Cirrus			Spectralis		
		AMD	*DME*	*RVO*	*AMD*	*DME*	*RVO*
IRC	Precision	0.71 (0.27)	0.76 (0.15)	0.72 (0.19)	0.78 (0.14)	0.78 (0.09)	0.85 (0.08)
	Recall	0.33 (0.22)	0.64 (0.17)	0.62 (0.17)	0.63 (0.13)	0.58 (0.14)	0.79 (0.10)
	n	56	16	68	50	16	10
SRF	Precision	0.82 (0.18)	0.84 (0.13)	0.87 (0.10)	0.81 (0.25)	0.90 (0.07)	0.89 (0.09)
	Recall	0.59 (0.27)	0.70 (0.21)	0.51 (0.30)	0.71 (0.24)	0.67 (0.29)	0.86 (0.11)
	n	63	8	8	45	10	10

AMD = age-related macular degeneration; DME = diabetic macular edema; IRC = intraretinal cystoid fluid; RVO = retinal vein occlusion; SRF = subretinal fluid.
Volume-based mean precision and recall over voxel-wise segmentations evaluated separately for IRC and SRF on neovascular AMD, DME, and RVO cases. The corresponding standard deviations are specified in parentheses.

not comprise IRC or SRF were not taken into account for the computation of the specific means for IRC and SRF.

We also evaluated the fluid quantification performance on the basis of the correlation of volume-wise automated and manual fluid volume measured as number of fluid pixels per volume. In these experiments, Pearson's correlation coefficient and R^2 were used as performance measures.

Results

Fully automated detection and quantification of macular fluid were possible in all included OCT scans. In the following, we report results on detection of the presence of intraretinal and SRF, as well as on full 3-dimensional quantification of fluid.

Detection of Fluid Presence

Both hallmark features of exudative disease activity, IRF and SRF, were detected with high accuracy in OCT images obtained by the 2 devices, Spectralis and Cirrus, in all 3 diseases, neovascular AMD, DME, and RVO. Even in DME in which SRF occurred less frequently and in smaller volumes, the results were found to be reliable. We evaluated the detection performance on 1200 OCT scans. Specifically, we used 200 scans per disease/device, that is, 100 scans with fluid and 100 without fluid per group. A summary of the results for detection of IRC and SRF per disease/device is provided in Table 1. The computations of performance (i.e., AUC) and recall were based on the ROC curves shown in Figure 2. Because precision is a measure of relevance of results and recall measures the proportion of relevant results, high precision attributes a classifier to yield accurate results, whereas a high recall means that the majority of all positive samples are truly detected.

Quantification of Fluid

Quantitative measurements of IRC and SRF within full OCT volumes were highly precise and reliable in all observed diseases and with the use of both devices using the new method. We computed pixel-wise segmentations of IRC and SRF for full volumes and evaluated the segmentation performance on the basis of volume-wise precision and recall values. The average volume-wise precision and recall per disease/device group are provided in Table 2.

Correlation of Clinical and Automated Measurements

The comparison with evaluation by expert readers trained for retinal analyses showed excellent correlation between automated and manual delineation of fluid features. We correlated the sizes of segmentations, measured as total number of pixels classified as IRC or SRF per OCT volume, with sizes of IRC or SRF of corresponding ground truth annotations. Correlation plots of the volume-wise segmented versus true fluid volumes are shown in Figure 3. Corresponding quantitative results are provided in Table 3.

Example cases of segmentation results on all evaluated disease/device groups in conjunction with corresponding raw images and ground truth annotations are illustrated in Figure 4. Original OCT images exhibit different classic conditions of fluid pooling in different layers with often indistinct demarcation in not preprocessed images explaining the difficulties and controversies in clinical diagnostics. Expert readers trained in standardized identification of fluid-related features perform fluid detection and delineation often superior to clinicians in clinical routine.[2] Fully automated segmentation offered identical precision in detection and delineation compared with the manual ground truth reading by experts in all 3 disease entities. Further example cases are provided as Videos 1–6 (available at www.aaojournal.org) showing entire OCT volumes.

Discussion

OCT imaging has become the mainstay of retinal diagnosis in clinical routine and in scientific studies. The management of the leading exudative macular diseases by vascular endothelial growth factor (VEGF) inhibition is largely based on the evaluation of retinal fluid for initial diagnosis and retreatment indications. However, a reliable and quantitative identification of fluid-related features is beyond the clinicians' capacity. We present a validated artificial intelligence method for a fully automated detection and quantification of macular fluid (i.e., IRC and SRF) in clinical OCT imaging of the 3 most prevalent retinal diseases, neovascular AMD, DME, and RVO. In terms of detecting the presence of fluid in routine OCT scans, our digital method achieved accuracies in the range of the inter-observer agreement between certified retinal experts reported in the literature.[13] On the basis of extensive validation in 1200

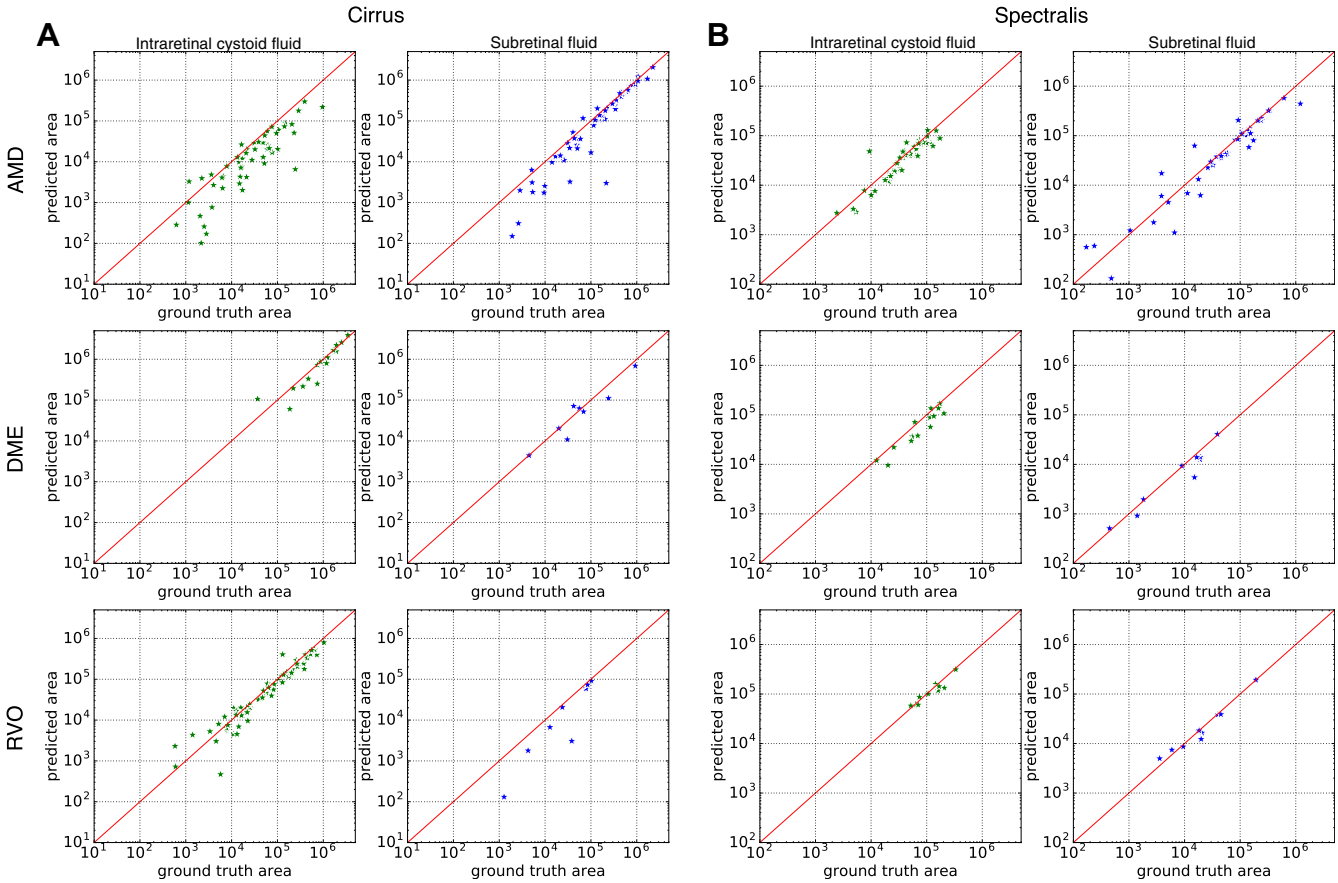

Figure 3. Correlation of automatically segmented versus manually assessed fluid volumes per OCT scan. **First row**: Neovascular age-related macular degeneration (AMD). **Second row**: Diabetic macular edema (DME). **Bottom row**: Retinal vein occlusion (RVO). **A**, Cirrus data. **B**, Spectralis data.

eyes with a broad range of exudative macular diseases and including Cirrus and Spectralis instruments, we believe that the presented method has reached a level of accuracy and reproducibility that may be acceptable for a wide application in clinical practice. Furthermore, we demonstrated that the proposed approach reliably and accurately quantifies fluid-filled regions and is capable of clearly distinguishing between IRC and SRF. Our method uses state-of-the-art artificial intelligence technology to solve the practical task of active fluid detection and quantification.

The presence of macular fluid represents the most important diagnostic retreatment criterion in the management of patients with exudative macular disease, and the evaluation of the fluid status on OCT has become a routine daily task not only for retina specialists but also for ophthalmologists in practice globally. On the basis of standard imaging, a reliable evaluation of leakage activity is often controversial, which matters particularly if retreatment is aimed at stabilization of disease activity such as in DME. A manual inspection of large OCT datasets in busy clinics is inherently impractical and prone to errors.[2] Moreover, modern OCT devices provide a huge number of images in the range of several hundreds, with swept-source OCT in the thousands as imaging technology continues to offer rapid improvements, providing more and more scans per imaging session together with high-resolution visualization of more and more morphologic features. Thus, automated methods present an attractive tool to at least prescreen OCT datasets and direct the clinician's attention to those images requiring detailed analysis, or in the near future perform independently. Our presented deep learning method provides solid accuracy for the detection of fluid, but, as opposed to prior work, is not limited to a particular disease or OCT device and may become widely applicable.

Detecting the mere presence of any fluid at any macular location is insufficient to guide clinical decision making, because both the type and the amount of fluid in the macula are important factors for patients' outcome. Prior studies, for example, in the setting of neovascular AMD indicate that approximately one quarter of eyes do not achieve a completely dry macula even after a prolonged period of intensive VEGF inhibition.[14] Nevertheless, some eyes may still be successfully extended under a treat-and-extend regimen in the sense that the amount of fluid does not increase when treatment is withheld. The value of so-called no tolerance regimens has been questioned by the scientific community.[15] Evidence suggests that not all fluid observed in eyes with neovascular AMD may be caused by active exudation, but may rather be a

Table 3. Fluid Volume Correlation ρ

Disease	Measure	Cirrus		Spectralis	
		IRC	*SRF*	*IRC*	*SRF*
AMD	ρ	0.81	0.98	0.86	0.85
	P	<0.00001	<0.00001	<0.00001	<0.00001
	R^2	0.39	0.94	0.68	0.65
	n	56	63	50	45
DME	ρ	0.98	0.99	0.88	0.96
	P	<0.00001	<0.00001	0.00001	0.00002
	R^2	0.92	0.89	0.62	0.87
	n	16	8	16	10
RVO	ρ	0.95	0.96	0.95	1.00
	P	<0.00001	0.00017	0.00003	<0.00001
	R^2	0.87	0.83	0.87	1.00
	n	68	8	10	10

AMD = age-related macular degeneration; DME = diabetic macular edema; IRC = intraretinal cystoid fluid; RVO = retinal vein occlusion; SRF = subretinal fluid.
Correlation of segmented and true fluid volume evaluated on Cirrus and Spectralis data, each grouped in neovascular AMD, DME, and RVO cases. For each result, Pearson's correlation coefficient ρ is specified, together with the 2-tailed *P* value and R^2. The evaluation was performed separately for IRC and SRF.

consequence of degenerative processes such as in "pseudocysts" and "outer retinal tabulation."[16,17] Particularly, persistent intraretinal cystoid changes are known to reflect neurosensory degeneration rather than active neovascular leakage, which does not resolve under therapy.[18] It is critical to obtain a precise measurement of the amount, increase, or decrease of fluid within or underneath the retina and use the dynamics of this quantitative variable to guide retreatment indications. A disciplined approach to retreatment strategies in respect to the socioeconomic burden of long-term disease activity is also warranted.[19] In DME, recent publications have shown that anti-VEGF therapy can "cure" DME after an initial period of active treatment, with many patients requiring no injections beyond the third year of therapy. Monitoring the precise amount of fluid, that is, leakage activity, is especially important when introducing a new retreatment regimen based on a "no change" policy. Deciding whether IRC has not changed is only feasible if IRC is automatically detected, the volume can be calculated, and the distribution of the IRC can be compared automatically.[20] Our method enables these approaches in a reliable, inexpensive, and time-saving manner.

Prior studies demonstrated a clear correlation between the amount of IRC at the foveal center and the best-corrected visual acuity.[6] Thus, a tool to quickly and reliably measure the amount and localization of fluid may become crucial to inform individual prognosis or as a manner to stratify patients in clinical trials. Recent structure/function correlation highlights the need to differentiate between the different types of macular fluid, that is, IRC and SRF.[21] Growing evidence suggests that these 2 types of fluid exert a different effect on visual acuity, with IRC being the main driver for functional loss and SRF conferring a possibly enhancing visual prognosis.[4] Indeed, eyes that did not show SRF at baseline had somewhat poorer functional outcomes of anti-VEGF therapy across the different macular diseases, including neovascular AMD and DME, also when balancing for baseline visual acuity levels.[4,22] More important, researchers have suggested that chronic, treatment-refractory SRF may not represent a vital indication for retreatment.[23] Although the prospective FLUID study (clinicaltrials.gov identifier, NCT01972789) that is currently under way will provide further evidence on these phenomena, it is clear that an automated method in clinical practice is extremely useful to detect and quantify IRC and SRF separately.[24]

Deep learning methods have been recently suggested for automated detection of diabetic retinopathy from fundus images.[25,26] Lee et al[27] used deep learning to link OCT images to clinical end points extracted from electronic medical records. The study neither aimed to detect nor to quantify neovascular AMD-related findings in OCT images, such as IRC, SRF, retinal pigment epithelium lesions, or drusen, but to distinguish normal OCT images from images of subjects affected by AMD. Detection and segmentation of retinal fluid have lately attracted substantial interest from several image analysis groups because currently no commercial OCT device offers such functionality. Intensive efforts in the development of automated algorithms highlight the attractivity of those tools for an advanced management of macular and retinal disease. Chen et al[28] used a graph-search-graph-cut approach for automated 3-dimensional segmentation of symptomatic exudate-associated derangement (SEAD), which subsumes IRC, SRF, and pigment epithelial detachment into a single fluid class, resulting in a binary classification between SEAD and non-SEAD. The method was evaluated on 15 OCT images of subjects with exudative AMD only annotated by a single grader on iPad software.[28] Fernández applied a deformable model to yield shape descriptions of fluid-filled regions on OCT scans of patients with AMD.[29] Wilkins et al[30] performed automated segmentation of IRC. The implemented method did not leverage deep learning methods but mainly relied on classic image processing methods, such as thresholding and boundary tracing. Their method was evaluated on Cirrus OCT scans of 16 patients with vitreoretinal disease and 3 controls only. Recently, efforts were aimed at achieving automated pixel-wise fluid segmentation. Xu et al[31] presented a machine learning—based approach where each pixel is independently classified using a handful of handcrafted features describing local texture. They evaluate their method on images from 10 patients with CNV, but the method does not differentiate IRC and SRF. Chiu et al[32] introduced a joint layer and fluid segmentation approach and evaluated it on 110 B-scans from 10 patients with DME. Although the method offers the potential to separate IRC from SRF based on layer information, the evaluated patients exhibited only IRC. Wang et al[33] presented a level-set—based fluid segmentation approach and evaluated it on 10 eyes with DME imaged with OCT angiography. Manual correction of segmented layers was required to separate IRC from SRF. In summary, previously reported methods focused on a single disease only (AMD or DME), were evaluated on a small number of patients, and did not differentiate automatically between IRC and SRF, but rather performed this separation in a postprocessing step. Our study goes far beyond previously published work in both performances, as well as the size and diversity of the validation dataset, which makes the automated methodology practically applicable.

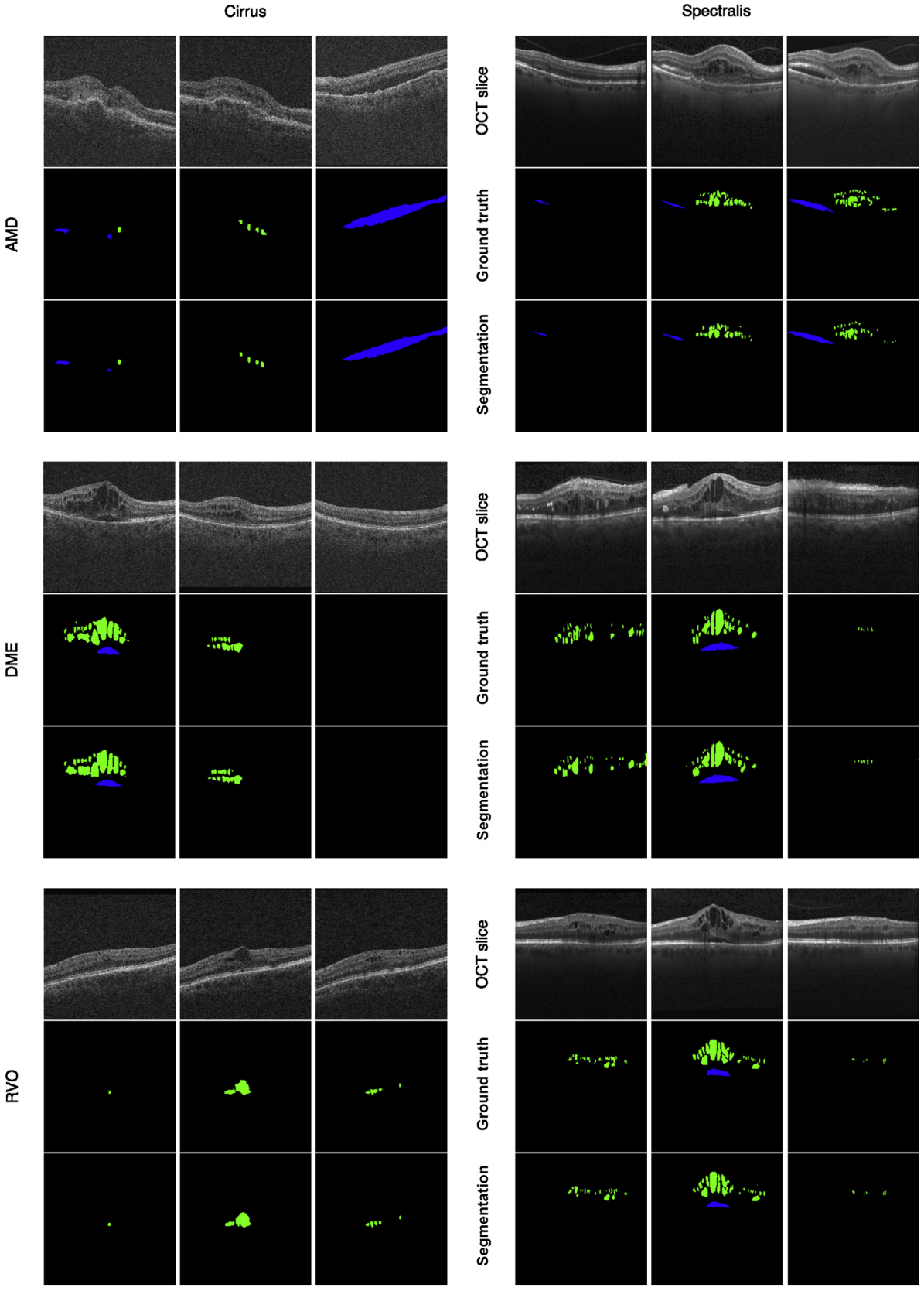

Figure 4. Example cases of segmentation results on OCT scans acquired with Cirrus or Spectralis devices of neovascular age-related macular degeneration (AMD), diabetic macular edema (DME), and retinal vein occlusion (RVO) cases. The **first row** shows OCT slices, the **second row** shows manual labels by certified graders, and the **bottom row** shows the automated segmentation results (intraretinal cystoid fluid [IRC] in *green*, SRF in *blue*).

Despite these efforts, fluid segmentation and detection represent a challenging problem because of the low signal-to-noise ratio of OCT images and the large variability in fluid shape and size, in particular across different retinal diseases.

In this article, we extend and evaluate the approach of Schlegl et al,[34] which introduced a deep learning approach for classifying pixels and segmenting IRC and SRF separately with a high level of accuracy, in 2 respects. First, in the aforementioned work, the model was trained and tested solely on Cirrus AMD data. Second, the method only was allowed to predict a single class label for a given input image patch. Besides intensity normalization of input images, the proposed method does not require any image preprocessing, thus avoiding error propagation typical of long analysis pipelines. It uses the most innovative artificial intelligence technology, termed "deep learning." The overwhelming wealth of clinical images and the distinct nature of morphologic features offer an ideal condition for deep learning methods to introduce groundbreaking paradigm shifts into the management of the leading retinal diseases of modern ophthalmology.

The detection performance for SRF in DME eyes indicates a potential for improvements. The poorer performance can be explained by the fact that, compared with neovascular AMD or RVO, DME eyes in general comprise a lower prevalence and lesser amount of SRF. Thus, when training the method solely on AMD and RVO cases and applying the trained model on DME cases, the model is not able to adapt for the different distributions of class labels. On the other hand, when the method would be directly trained on DME cases, because of the small amount of target labels, it would not be able to acquire a good representation for SRF. Because of differences in label distributions, simultaneous training on AMD, RVO, and DME cases would reduce the test performance on DME. This limitation should be overcome by the availability of sufficient amounts of ground truth cases.[35] Future deep learning methods may not require large amounts of ground truth annotations any longer because they could be designed to detect deviations from normal-appearance patterns in an unbiased manner.

In conclusion, we have presented and validated a highly robust and sensitive automated method to detect, differentiate, and quantify macular fluid (IRC and SRF) in routine OCT images of a range of exudative macular diseases as well as the most widely used OCT devices. Because of their precision, reliability, and objectivity, digital analysis tools such as the one presented may become important tools in the individual and the large-scale management of patients with macular disease. Deep learning presents a promising technology for reliably guiding disease and patient management in a field so intensively driven by imaging such as ophthalmology.

References

1. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254:1178−1181.
2. Toth CA, Decroos FC, Ying GS, et al. Identification of fluid on optical coherence tomography by treating ophthalmologists versus a reading center in the Comparison of Age-Related Macular Degeneration Treatments Trials. *Retina*. 2015;35:1303−1314.
3. Schmidt-Erfurth U, Waldstein SM. A paradigm shift in imaging biomarkers in neovascular age-related macular degeneration. *Prog Retin Eye Res*. 2016;50:1−24.
4. Waldstein SM, Simader C, Staurenghi G, et al. Morphology and visual acuity in aflibercept and ranibizumab therapy for neovascular age-related macular degeneration in the VIEW Trials. *Ophthalmology*. 2016;123:1521−1529.
5. Chakravarthy U, Goldenberg D, Young G, et al. Automated identification of lesion activity in neovascular age-related macular degeneration. *Ophthalmology*. 2016;123:1731−1736.
6. Waldstein SM, Philip AM, Leitner R, et al. Correlation of 3-dimensionally quantified intraretinal and subretinal fluid with visual acuity in neovascular age-related macular degeneration. *JAMA Ophthalmol*. 2016;134:182−190.
7. Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. Proceedings of the IEEE International Conference on Computer Vision; Santiago de Chile, Chile, December 7 − 13, 2015.
8. Chetlur S, Woolley C, Vandermersch P, et al. cuDNN: Efficient Primitives for Deep Learning. Cornell University Library: https://arxiv.org/abs/1410.0759 (submitted on 3Oct2014). arXiv preprint arXiv:14100759 2014.
9. James B, Olivier B, Frédéric B, et al. Theano: a CPU and GPU math expression compiler. Proceedings of the Python for Scientific Computing Conference (SciPy). https://conference.scipy.org/proceedings/scipy2010/bergstra.html.
10. Dieleman S, Schlüter J, Raffel C, et al. *Lasagne: First Release*. Geneva, Switzerland: Zenodo; 2015.
11. Garvin MK, Abramoff 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.
12. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32−35.
13. Wu J, Philip AM, Podkowinski D, et al. Multivendor spectral-domain optical coherence tomography dataset, observer annotation performance evaluation, and standardized evaluation framework for intraretinal cystoid fluid segmentation. *J Ophthalmol*. 2016;2016:3898750.
14. Maguire MG, Martin DF, Ying GS, et al. Five-year outcomes with anti-vascular endothelial growth factor treatment of neovascular age-related macular degeneration: the comparison of age-related macular degeneration treatments trials. *Ophthalmology*. 2016;123:1751−1761.
15. Freund KB, Korobelnik J-F, Devenyi R, et al. Treat-and-extend regimens with anti-VEGF agents in retinal diseases: a literature review and consensus recommendations. *Retina*. 2015;35:1489−1506.
16. Querques G, Coscas F, Forte R, et al. Cystoid macular degeneration in exudative age-related macular degeneration. *Am J Ophthalmol*. 2011;152:100−107.e2.
17. Giachetti Filho RG, Zacharias LC, Monteiro TV, et al. Prevalence of outer retinal tubulation in eyes with choroidal neovascularization. *Int J Retina Vitreous*. 2016;2:6.
18. Schmidt-Erfurth U, Waldstein SM, Deak GG, et al. Pigment epithelial detachment followed by retinal cystoid degeneration leads to vision loss in treatment of neovascular age-related macular degeneration. *Ophthalmology*. 2015;122:822−832.
19. Gupta OP, Shienbaum G, Patel AH, et al. A treat and extend regimen using ranibizumab for neovascular age-related macular degeneration clinical and economic impact. *Ophthalmology*. 2010;117:2134−2140.
20. Elman MJ, Ayala A, Bressler NM, et al. Intravitreal ranibizumab for diabetic macular edema with prompt versus

deferred laser treatment: 5-year randomized trial results. *Ophthalmology*. 2015;122:375–381.
21. Sharma S, Toth CA, Daniel E, et al. Macular morphology and visual acuity in the second year of the Comparison of Age-Related Macular Degeneration Treatments Trials. *Ophthalmology*. 2016;123:865–875.
22. Sophie R, Lu N, Campochiaro PA. Predictors of functional and anatomic outcomes in patients with diabetic macular edema treated with ranibizumab. *Ophthalmology*. 2015;122:1395–1401.
23. Arnold JJ, Markey CM, Kurstjens NP, Guymer RH. The role of sub-retinal fluid in determining treatment outcomes in patients with neovascular age-related macular degeneration—a phase IV randomised clinical trial with ranibizumab: the FLUID study. *BMC Ophthalmol*. 2016;16:31.
24. Schmidt-Erfurth U, Klimscha S, Waldstein SM, Bogunovic H. A view of the current and future role of optical coherence tomography in the management of age-related macular degeneration. *Eye (Lond)*. 2017;31:26–44.
25. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*. 2017;124:962–969.
26. Abramoff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci*. 2016;57:5200–5206.
27. Lee CS, Baughman DM, Lee AY. Deep learning is effective for classifying normal versus age-related macular degeneration optical coherence tomography images. *Ophthalmol Retina*. 2017;1:322–327.
28. Chen X, Zhang L, Sohn EH, et al. Quantification of external limiting membrane disruption caused by diabetic maular edema from SD-OCT. *Invest Ophthalmol Vis Sci*. 2012;53(13):8042–8048.
29. Fernandez DC. Delineating fluid-filled region boundaries in optical coherence tomography images of the retina. *IEEE Trans Med Imaging*. 2005;24:929–945.
30. Wilkins GR, Houghton OM, Oldenburg AL. Automated segmentation of intraretinal cystoid fluid in optical coherence tomography. *IEEE Trans Biomed Eng*. 2012;59:1109–1114.
31. Xu X, Lee K, Zhang L, et al. Stratified sampling voxel classification for segmentation of intraretinal and subretinal fluid in longitudinal clinical OCT data. *IEEE Trans Med Imaging*. 2015;34(7):1616–1623.
32. Chiu SJ, Allingham MJ, Mettu PS, et al. Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. *Biomed Opt Express*. 2015;6:1172–1194.
33. Wang J, Zhang M, Pechauer AD, et al. Automated volumetric segmentation of retinal fluid on optical coherence tomography. *Biomed Opt Express*. 2016;7:1577–1589.
34. Schlegl T, Waldstein SM, Schmidt-Erfurth U, Langs G. predicting semantic descriptions from medical images with convolutional neural networks. *Inf Process Med Imaging*. 2015;24:437–448.
35. Gerendas B, Simader C, Deak GG, et al. Morphological parameters relevant for visual and anatomic outcomes during anti-VEGF therapy of diabetic macular edema in the RESTORE trial. *Invest Ophthalmol Vis Sci*. 2014;55:1791.

Footnotes and Financial Disclosures

Originally received: May 4, 2017.
Final revision: October 23, 2017.
Accepted: October 24, 2017.
Available online: December 8, 2017. Manuscript no. 2017-1041.

[1] Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University Vienna, Vienna, Austria.

[2] Department of Ophthalmology, Medical University of Vienna, Vienna, Austria.

[3] Vienna Reading Center, Department of Ophthalmology, Medical University of Vienna, Vienna, Austria.

[4] Computational Imaging Research Lab Department of Biomedical Imaging and Image-Guided Therapy, Medical University Vienna, Vienna, Austria.

Financial Disclosure(s):
The author(s) have made the following disclosure(s): T.S.: Patent pending – WO 2016139183 A1.

S.M.W.: Consultancy – Bayer, Novartis; Research support – Bayer, Genentech; Patent pending – WO 2016139183 A1.

T.S., S.M.W., H.B., F.E., A.S., A.-M.P., D.P., B.S.G., G.L., and U.S.-E.: Grants – Christian Doppler Research Association.

A.-M.P.: Grants – Austrian Federal Ministry of Economy, Family and Youth, National Foundation of Research, Technology and Development.

B.S.G.: Consultancy – Roche; Patent pending – WO 2016139183 A1.

G.L.: Patent pending – WO 2016139183 A1.

U.S.-E.: Consultancy – Bayer, Boehringer Ingelheim, Novartis, Roche; Patent pending – WO 2016139183 A1.

This study received funding from IBM (TS: 2016-2017 IBM PhD Fellowship Award), the Austrian Federal Ministry of Science, Research and Economy, Novartis, and the Austrian Science Fund (I 2714-B31). The sponsor or funding organizations had no role in the design or conduct of this research.

HUMAN SUBJECTS: Human subjects were included in this study. All studies adhered to the tenets of the Declaration of Helsinki. Study protocol was approved by the Ethics Committee of the Medical University of Vienna.

Author Contributions:

Conception and design: Schlegl, Waldstein, Bogunovic, Schmidt-Erfurth

Data collection: Schlegl, Endstraßer, Sadeghipour, Philip, Podkowinski, Gerendas

Analysis and interpretation: Schlegl, Waldstein, Bogunovic, Langs, Schmidt-Erfurth

Overall responsibility: Schlegl, Waldstein, Bogunovic, Sadeghipour, Gerendas, Langs, Schmidt-Erfurth

Abbreviations and Acronyms:
AMD = age-related macular degeneration; **AUC** = area under the curve; **DME** = diabetic macular edema; **IRC** = intraretinal cystoid fluid; **ROC** = receiver operating characteristic; **RVO** = retinal vein occlusion; **SEAD** = symptomatic exudate-associated derangement; **SRF** = subretinal fluid; **VEGF** = vascular endothelial growth factor.

Correspondence:
Ursula Schmidt-Erfurth, MD, Professor and Chair, Department of Ophthalmology, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria. E-mail: ursula.schmidt-erfurth@meduniwien.ac.at.