Ultra-Widefield OCT Angiography

M. Niederleithner, L. de Sisternes, H. Stino, A. Sedova, T. Schlegl, H. Bagherinia, A. Britten, P. Matten, U. Schmidt-Erfurth, A. Pollreisz, W. Drexler, R. A. Leitgeb, and T. Schmoll

Abstract—Optical Coherence Tomography Angiography (OCTA), a functional extension of OCT, has the potential to replace most invasive fluorescein angiography (FA) exams in ophthalmology. So far, OCTA's field of view is however still lacking behind fluorescence fundus photography techniques. This is problematic, because many retinal diseases manifest at an early stage by changes of the peripheral retinal capillary network. It is therefore desirable to expand OCTA's field of view to match that of ultra-widefield fundus cameras. We present a custom developed clinical high-speed swept-source OCT (SS-OCT) system operating at an acquisition rate 8-16 times faster than today's state-of-the-art commercially available OCTA devices. Its speed allows us to capture ultra-wide fields of view of up to 90 degrees with an unprecedented sampling density and hence extraordinary resolution by merging two single shot scans with 60 degrees in diameter. To further enhance the visual appearance of the angiograms, we developed for the first time a three-dimensional deep learning based algorithm for denoising volumetric OCTA data sets. We showcase its imaging performance and clinical usability by presenting images of patients suffering from diabetic retinopathy.

Index Terms—OCT, OCTA, SS-OCT, widefield, angiography, ophthalmic imaging.

M. Niederleithner, T. Schlegl, A. Britten, P. Matten, W. Drexler, and R. A. Leitgeb are with the Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, 1090 Vienna, Austria (e-mail: michael.niederleithner@meduniwien.ac.at; thomas.schlegl@meduniwien.ac.at; anja.britten@meduniwien.ac.at; philipp.matten@meduniwien.ac.at; wolfgang.drexler@meduniwien.ac.at; rainer.leitgeb@meduniwien.ac.at).

L. de Sisternes and H. Bagherinia are with Carl Zeiss Meditec Inc., Dublin, CA 94568 USA (e-mail: luis.desisternes@zeiss.com; homayoun.bagherinia@zeiss.com).

H. Stino, A. Sedova, U. Schmidt-Erfurth, and A. Pollreisz are with the Department of Ophthalmology, Medical University of Vienna, 1090 Vienna, Austria (e-mail: heiko.stino@meduniwien.ac.at; aleksandra.sedova@meduniwien.ac.at; ursula.schmidt-erfurth@meduniwien.ac.at; andreas.pollreisz@meduniwien.ac.at).

T. Schmoll is with Carl Zeiss Meditec Inc., Dublin, CA 94568 USA, and also with the Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, 1090 Vienna, Austria (e-mail: tilman.schmoll@zeiss.com).

This article has supplementary downloadable material available at https://doi.org/10.1109/TMI.2022.3222638, provided by the authors.

Digital Object Identifier 10.1109/TMI.2022.3222638

I. INTRODUCTION

THE retina with its fine capillary network is one of the most perfused tissues of the human body. Retinal diseases, such as age-related macular degeneration, diabetic retinopathy or glaucoma lead to pathologic changes of this complex network. There is evidence that many ophthalmic diseases affect the capillary network in their early stages especially in the retinal periphery [1], [2], [3], [4], [5].

In general, an angiogram is a map of the vascular network. Modern fundus cameras provide angiograms with large fields of view (FOVs) of 90 degrees and more using fluorescent contrast agents, fluorescein or indocyanine green. Those agents need to be injected, making the procedure invasive and time consuming. Further, the finest capillaries of the retina cannot be resolved in widefield fundus angiographies (FA) [6], [7], [8]. On the other hand, optical coherence tomography angiography (OCTA) [9], [10], [11], [12], [13], [14], a functional extension of optical coherence tomography (OCT), can create depth resolved maps of the retinal vascular network without the administration of a contrast agent and with the ability to resolve even the finest capillaries.

OCT creates volumetric structural data sets based on interferometric information of partial coherent light [15], [16], [17], [18], [19], [20]. Motion sensitive extensions of OCT, such as Doppler OCT and OCTA, have been shown to contrast the retinal vasculature and to even provide quantitative blood flow information [21], [22], [24]. For visualization of blood vessels, OCTA measures the signal decorrelation between multiple repeated OCT acquisitions at the same location [10], [11], [12], [13], [25], [26]. This need for oversampling, increases the acquisition time or limits the FOV with respect to regular OCT scans. However, in order to replace most FA exams by OCTA scans, OCTA needs to reach comparable FOVs. Initially, the maximum achievable FOV of OCTA was limited to small patches of 3 mm × 3 mm and 6 mm × 6 mm. Angiograms of the macula and the optic nerve head therefore had to be acquired in separate scans. Today's commercially available devices enable OCTA images with a FOV of up to 15 mm × 15 mm, providing images including both macula and optic nerve head in one single acquisition [27]. Those "widefield" angiography images are however only sparsely sampled, resulting in loss of vascular detail especially when compared to 3 mm × 3 mm or 6 mm × 6 mm scans. To extend the area which can be examined, stitching can be used to create an assembly of several OCTA images acquired at different positions, covering a larger FOV in total [26].
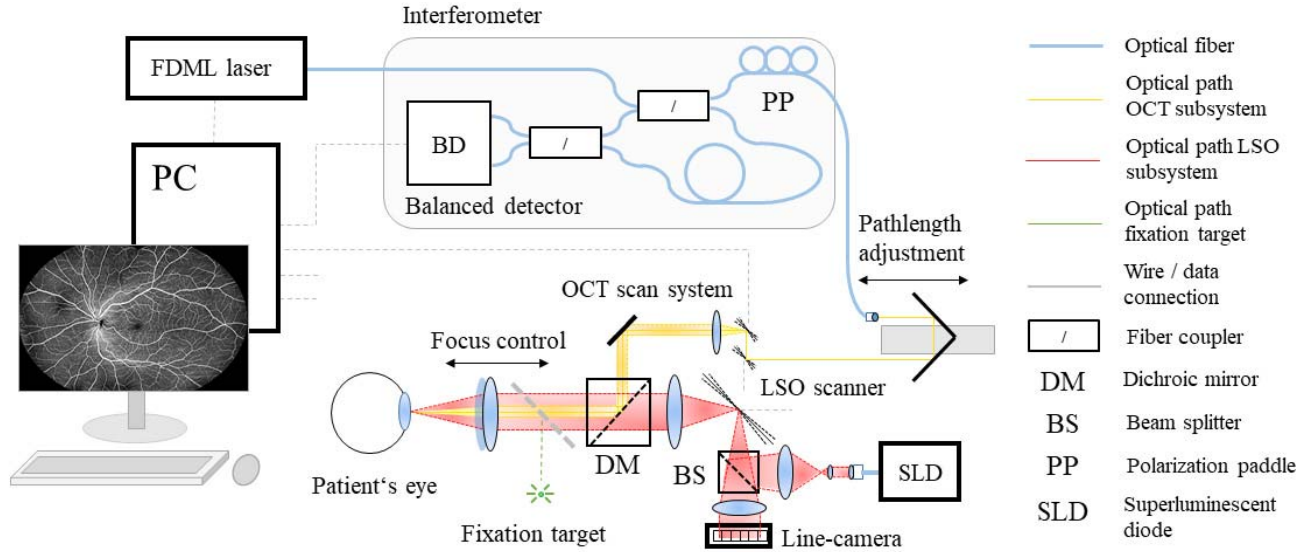
Fig. 1. Schematics of optical system with OCT subsystem and LSO subsystem used for retinal motion correction.

Modern commercially available OCTA devices offer such stitching capabilities, providing a total FOV up to 21 mm × 21 mm [27]. However, these imaging modes are rarely used in clinical practice. They require several individual scans with fixation points far in the periphery, making them inconvenient for the patient and difficult to acquire with good image quality. Stitching multiple of such acquisitions is also prone to artifacts due to image distortions or differences in image quality between scans. Because stitching is so far done using en face images, OCTA's volumetric information is in part lost and a retrospective analysis of custom slabs is in these cases typically not possible.

The FOV for a given acquisition time is set by the lateral sampling density and the lateral sampling rate, i.e., the system's A-scan rate. Because the lateral sampling cannot be arbitrarily decreased without losing vascular details, we chose to significantly increase the A-scan rate compared to previously published OCTA systems and state of the art commercial OCTA systems, which still operate at a few hundred kHz A-scan rate. However, increasing A-scan rate and FOV while keeping the acquisition time constant, comes at the cost of reduced SNR and thereby also increased phase noise, making OCTA with high image quality more challenging [28].

Fourier-domain-mode-locking (FDML) lasers are swept-source lasers, which offer sweep-rates up to several MHz [29], [30]. Systems equipped with such lasers have previously been shown to enable large FOVs. Kolb et al. showed widefield structural OCT with a FOV of 100 degrees, but without OCTA [31], [32]. Blatter et al. used a FDML laser source operating at 1.68 MHz A-scan rate to capture OCTA images of 12 mm × 12 mm [9]. Although the sensitivity was reduced compared to slower systems, at least larger capillaries could be resolved with good contrast over the full FOV. A comparison of typical A-scan rates and sensitivities is provided in Table I.

TABLE I
A-SCAN RATES OF TODAY'S COMMERCIAL OCTA SYSTEMS

SYSTEM	A-scan rate [kHz]	Sensitivity [dB]
Spectral domain	60 – 100	> 105
Swept source	100 – 200	105
Prototype	1,680	95

Recently, Wei et al. used a 400 kHz swept-source to create OCTA images with a claimed FOV of 75 degrees [33], [34], but with a relatively long acquisition time of approximately one minute using a motion detection algorithm and rescanning clusters where motion occurred. Deep learning methods have been shown to denoise OCTA en face projections [35] or perform the whole OCTA-processing on the B-scan level [36], [37]. Such algorithms can improve image quality drastically, allowing us to compensate for the lower SNR resulting from the increased A-scan rate. To utilize the full volumetric information, we developed for the first time a 3-dimensional denoising algorithm, based on a U-net structure and working on volumetric OCTA data.

Using our high-speed SS-OCT device paired with this denoising algorithm, we demonstrate ultra-widefield clinical OCTA scans with significantly better image quality than state of the art OCTA scans acquired with much slower commercial OCTA systems.

II. METHODS

A. Imaging System

A schematic of the developed prototype can be seen in Fig. 1. As a laser source we used a Frequency-Domain-Mode-Locked (FDML) laser operating at a sweep-rate of

1.68 MHz (Optores GmbH, Munich, Germany). Unlike most swept-source lasers, the laser uses a fiber ring buffer with a Fabry-Perot filter instead of a traditional resonator. The roundtrip time of the ring buffer is matched to the frequency of the filter of 419 kHz to be able to buffer a full sweep at a time. Therefore, lasing does not have to be build up from the spontaneous background for every single optical frequency as is the case in commonly used external cavity swept lasers (ECTLs). To further increase the base sweep-rate of 419 kHz, the central, fairly linear part of the up-sweep is selected and buffered twice in an optical fiber buffer external to the actual laser cavity. This results in 1.68 MHz sweep-rate and 75 nm maximum tuning range with 100% duty cycle at a center wavelength of 1060 nm [29]. Those spectral characteristics result in a measured axial resolution of 9 μm (FWHM) in tissue.

The optical system consists of two subsystems, which are merged using a dichroic mirror: a confocal point-scanning OCT system and a line scanning ophthalmoscope (LSO) for observation of the patient's fundus as well as for tracking eye motion. The delivery optics to the eye were all custom designed. The OCT subsystem uses a Mach-Zehnder type interferometer, where the reference path is a closed optical fiber and the path length adjustment is done in the free space portion of the sample arm. With that, the light does only have to be coupled out once in the sample arm and the whole interferometer stays compact and simple. To match the polarization of sample and reference light, a motorized polarization paddle is located in the sample arm. The scan system consists of a pair of galvanometer scanners (6215H with high power option, Cambridge Technologies, Bedford, MA, USA) and provides a scan angle of 60 degrees, translating to a FOV of 18 mm in diameter on the retina. The field of view was verified with a model eye exhibiting a checkerboard pattern on the fundus, with 1 mm × 1 mm cell size spanning 10 mm × 10 mm. A custom designed optical relay in 4f-configuration delivers a scanned beam of 4.5 mW incident on the patient's pupil at a working distance of 27 mm. Because the scan-system as well as laser power is permanently monitored and the light source is shut down immediately in case of a failure, this power level is safe according to laser safety and ophthalmic device standards [38], [39], [40]. The beam diameter is 1 mm at the cornea, corresponding to a spot size of approx. 20 μm ($1/e^2$) on the retina. For the OCT subsystem we measured a sensitivity of 95 dB with a −6 dB roll-off at 3 mm depth in air.

Due to the retinal curvature an imaging depth of several millimeters is necessary for widefield OCT. We use a dual balanced detector with a bandwidth of 1.6 GHz (PDB480C-AC, Thorlabs, Newton, NJ, USA) and a digitizer board with a maximum sampling rate of 4 GS/s (ATS 9373, Alazartech Inc., Pointe Claire, Canada) to enable a dense spectral sampling and thereby large imaging depth. For this work, we sampled at a rate of approximately 2.5 GS/s given by a linear in time reference clock provided by the laser, yielding 1536 points per A-scan and an imaging depth of 6 mm in tissue. After acquisition the spectra are resampled to a linear sampling in wavenumber space. For that, a calibration measurement using two single reflectors was taken. Due to the high sweep repeatability of the FDML laser, this calibration measurement had only to be done once. The phase along the interference signal was extracted using a Hilbert transform and used as the resampling function [42].

The LSO uses a superluminescent diode with a central wavelength of 750 nm, delivering 1.1 mW on the pupil. The beam is formed into a line pattern with a FOV of 36 degrees, which is scanned across the retina to create a fundus image. The LSO (750 nm) is coupled into the OCT's (1060 nm) sample arm using a dichroic mirror. It operates with an enface update rate of 50 Hz. The successive images of the LSO are analyzed for motion and the lateral offsets are extracted. Small offsets of the patient's fixation are corrected continuously in the OCT scan-system. Areas where strong motion occurred will be rescanned. This results in volumetric data sets without lateral motion artifacts.

Due to the line illumination and the used power of the LSO subsystem it is save to be operated at the same time as the OCT subsystem according to the relevant safety standards [38], [39], [40].

For stabilization of the patient's gaze a fixation target is coupled into the sample arm using a second dichroic mirror.

To extend the FOV further into the periphery we also have the ability to merge multiple acquisitions with offset fixation positions in post-processing. We acquire 2 sub-images with a FOV of 18 mm in diameter each, where the fixation target only has to be moved by −15 degrees and +15 degrees horizontally from the center. This shortens the total acquisition time and improves the success rate significantly compared to multi-acquisition widefield OCTA scans on state-of-the-art commercial OCTA systems. These systems typically use 5 or more individual scans and more extreme fixation locations. Our resulting merged scans cover a horizontal FOV of approximately 90 degree or 27 mm and 60 degree or 18 mm vertically. Each acquisition takes less than 20 seconds. Because the fixation points are still conveniently located for the patient, realignment between the two individual scans is fast and the overall acquisition time is typically still below 1 minute per eye.

All patient data presented herein were acquired within a study reviewed by the ethics committee of the Medical University of Vienna and informed consent was obtained. The imaging procedure is performed in the following steps: The patient takes place in front of the device and is asked to look into the center of the fixation target. Then s/he will be aligned with a motorized head rest and the help of an iris camera. The polarization of the sample arm light and the zero-delay position are adjusted while the retina can be already observed with an OCT-preview. A live display of the co-aligned LSO is used to set the focus and to confirm that the retina is illuminated homogenously. Once the patient is in position, the acquisition can be started. During acquisition, the retina can be continuously observed via LSO and OCT-preview and the zero-delay position can be re-adjusted if needed. Because of the system's retinal tracker, the patient can blink, and the acquisition will continue as soon as the OCT and fundus image returns.

TABLE II
USED SCAN PATTERNS

TRAINING VOLUME	Parameter
Lateral points	1024 x 1024 (512)
Repetitions	8
FOV	9 mm x 9 mm (4.5 mm)
Pixel spacing	8.7 μm
B-scan rate	~ 1 kHz
Acquisition time	min 8.5 seconds

WIDEFIELD VOLUME	Parameter
Lateral points	2048 x 2048
Repetitions	2 or 4
FOV	18 mm x 18 mm (scan area)
Pixel spacing	8.7 μm
B-scan rate	~ 500 Hz
Acquisition time	min. 8 or 15 seconds

The retina is scanned in a traditional raster scan pattern with a horizontal fast axis. Due to the high A-scan rate of 1.68 MHz and the large scan angle, a flyback with 50% of the duration of the B-scan ramp is chosen to reduce stress on the fast axis scanner, resulting in a fast axis scan duty cycle of 75%. To work at such high B-scan rates the shape of the flyback function of the scanners was optimized to avoid abrupt accelerations of the scan mirrors. All acquisitions are sampled isotopically with a pixel spacing of 8.7 μm. Detailed parameters of the used scan patterns are listed in Table II.

To detect and correct for lateral motion, the images of the LSO-subsystem are analyzed on the fly. A feature detection algorithm is used for determining the shift relative to the previous image and a reference image, taken at the beginning of the acquisition. These shifts are used to correct the scan angle of the OCT subsystem. Further, if too large motion is detected, the corresponding clusters of B-scans are enqueued for rescanning. The correction of the scan angle is done right after the shifts are calculated, while the rescanning algorithm works dynamically during acquisition. The performance of the tracking algorithm might be further improved in accuracy, as can be seen by slight motion artifacts, e.g. in Fig. 7a. Currently, it can reliably correct offsets larger than approximately 50 μm, being sufficient for widefield imaging.

All patients were first imaged with a commercially available system PLEX® Elite 9000 (ZEISS, Dublin, CA). If this was successful, meaning we were able to reveal diagnostically relevant details, they were imaged with the presented prototype. Therefore, patients very difficult to image on the commercial device were automatically excluded. However, all patients who could be imaged with the commercial system, could also be imaged with the prototype.

In our clinical trail the patient with the highest myopia was −6 diopters and could be imaged comfortably. A video showcasing the alignment and acquisition procedure is attached as a Multimedia-File (Supplementary material: Patient Acquisition.mp4).

B. OCT and OCTA Image Reconstruction

For OCTA processing, a cluster of n B-scans at each slow scanning position is used. The first one of every cluster of such repeated B-scans is taken as a reference frame and the following are bulk motion corrected in axial direction using a cross-correlation-based algorithm, where the position of the maximum of the cross-correlation with respect to the center determines the relative axial shift to be corrected (Equation 1).

Given a shift of dz in the axial direction, the cross correlation between the first and the i-th B-scan of a cluster is defined as follows:

$$xcorr_{i,dz} = \sum_{x,y,z} |s_{x,y,z,1}| * |s_{x,y,(z+dz),i}| \tag{1}$$

where s is a complex voxel in space specified by depth position z (sample in an A-scan), fast axis position x (A-scan in a B-scan) and slow axis position y (B-scan in a volume).

Afterwards, the corresponding A-scans are sub-pixel motion corrected by subtracting the mean phase-difference between A-scans of the reference B-scan 1 and the corresponding ones of the other B-scans of a cluster (Equation 2), where Z is the total number of samples per A-scan.

$$\widehat{ph}_{x,y,i} = \frac{1}{Z} \sum_{z}^{Z} \frac{s_{x,y,z,i} * s^*_{x,y,z,(i-1)}}{|s_{x,y,z,i} * s^*_{x,y,z,(i-1)}|} \tag{2}$$

$$s_{x,y,z,i} = s_{x,y,z,i} * \hat{ph}_{x,y,i} \tag{3}$$

Then the flow signal is calculated as the complex variance between neighboring B-scans:

$$f_{x,y,z} = log_{10}(\frac{1}{n-1} \sum_{i=1}^{n-1} |s_{x,y,z,i} - s_{x,y,z,(i+1)}|^2) \tag{4}$$

where $f_{x,y,z}$ is the flow-voxel at position (x, y, z), n is the number of the B-scan within one cluster and s is a complex voxel. This results in n-1 sub-images per cluster having the same interscan time. These sub-images are averaged to create the final noise reduced OCTA B-scan for each slow-axis scanning position. Since OCT has a high dynamic range, log compression was used to gain more contrast of small flow-values.

All intensity B-scans of a cluster are averaged as well to improve SNR and are used for layer segmentation later on. These processing steps are repeated for each cluster at every slow axis scanning position.

Although OCTA produces volumetric data sets, angiograms are typically displayed as en face images. For an intuitive en face presentation of the vascular network at a particular sample depth it is of advantage to follow the natural shape of the sample. In our case we chose the RPE as a reference for the retinal shape. It offers a strong signal in the intensity B-scans which is well suited for segmentation. Because segmentation errors lead to artifacts in the en face projections, the algorithm has to be very robust, especially when imaging very large FOVs and patients with pathologically altered retinas.

Fig. 2. Schematics of segmentation steps.
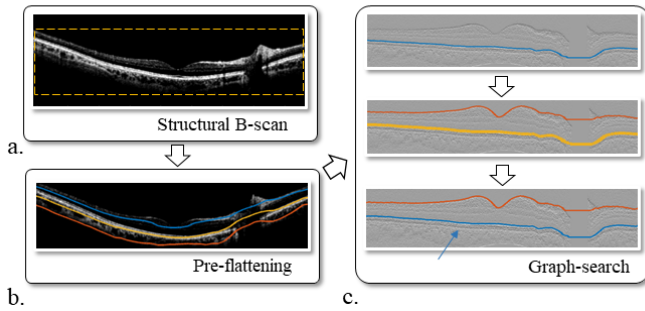

Fig. 3. Model used for denoising; given numbers are numbers filters.

To minimize segmentation errors around the optic nerve head (ONH) it is excluded from the cross-sectional segmentation algorithm. To do so, it is automatically segmented on the enface plane by a deep learning algorithm based on a U-net, described in [42]. After automated segmentation, all voxels in the full z-range located within the predicted ONH enface projection are excluded from further processing. Next, cross-sectional images are loaded and processed for each depth. Each intensity image is thresholded to exclude residual noise, smoothed with a Gaussian filter and cropped to a rectangular area of non-zero pixels (Figure 2a). Assuming the retinal pigment epithelium (RPE) is a continuous layer on a curved plane, the retina is first coarsely flattened to improve the robustness of the following graph-search algorithm. To do so, the depth position of each A-scan is calculated by its center of mass, i.e. its intensity weighted depth position. Using this information, each A-scan is shifted to a common depth to flatten the retina (Figure 2b, yellow line).

The layer of highest reflectivity can vary between RPE and the inner limiting membrane (ILM) due to imperfect patient alignment or changing focus. In some instances, the graph search segmentation may therefore jump between those two boundaries. To correct for this, each image is segmented twice. The weights are calculated according to Chiu et al. [43]. After the first iteration (blue line) the resulting segmentation line is masked for the second iteration by setting the weights along this line to 10^9(yellow line), preventing the algorithm from segmenting along masked areas (Figure 2c). The deeper segmentation line is determined to be the RPE and is used for another, more precise flattening step of the structural and flow volumes.

After segmentation, the volumetric data is smoothed with a 3-dimensional Gaussian filter. The en face projections are created using a maximum intensity projection along depth from ILM to the photoreceptor inner and outer segment junction, which was determined by a constant offset from the RPE. This visualizes the vasculature of the complete inner retina. Finally, a bandpass filter is applied to the en face images for attenuating stripe artifacts.

C. Three-Dimensional Denoising

Our model uses a simple 3-dimensional U-net architecture, shown in Figure 3. It consists of 3 down-sampling layers, followed by a bottleneck-layer and 3 up-sampling layers.
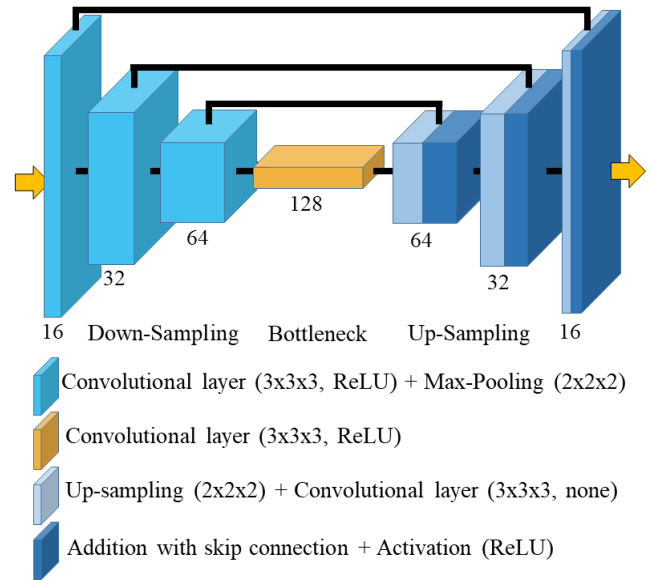
As being the key element of a U-net, the output of every of down-sampling layer is merged with the inputs of the corresponding up-sampling layers. The kernel-size of the convolutional layers is $3 \times 3 \times 3$ and the number of filters $32 \times i$, with i as the layer's index. All layers, except the output layer, use the rectified linear unit activation function (ReLU). For the output layer no activation was used. The used model contains 301,875 trainable parameters. It is trained by mapping low quality volumetric OCTA patches onto high quality patches. To generate the pairs of datasets we acquired OCTA volumes with 8 repetitions per slow axis position. The used scan pattern is described in Table II.

The high-quality data is used as the target, where all 8 repetitions are used for processing. To create the low-quality input data, the same datasets are used, but only the first 2 repetitions of a cluster are processed and the remaining 6 are discarded. Being based on the decorrelation between OCT scans, 2 is the minimum number needed for OCTA processing. Therefore, this is also the number of repetitions for the later acquired widefield images. The generation of the training data is depicted in Figure 4.

For this work only 9 volumes of healthy retinas with $128 \times 1024 \times 1024$ voxels are used, resulting in approximately 45,000 totally independent patches.

The interscan time is different for the training and the later used inference data. But our observations confirmed the findings of Choi et al. [44], that we do not lose signal from small capillaries by shortening the interscan time from 2 ms to 1 ms

Since a patch is selected with a random coordinate in space (x, y, z) the effective number of independent patches is much higher, but no further augmentation was used. The loss function was chosen to be the mean squared error with a small bias for numerical stability:

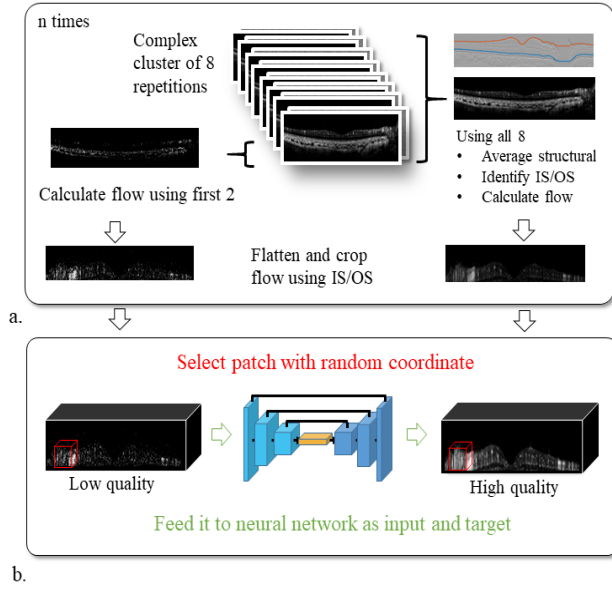$$L = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i + 10^{-8})^2 \tag{5}$$

Fig. 4.　Generation of training data and training process.

with x as the input, y the target and i as the voxel index. Adam was used as optimizer with a learning rate of 10^{-4}.

Training was performed using Keras (version 2.2.4) on a GeForce GTX 1060 GPU (nVidia).

In order to prevent overfitting on a single training set, the deep learning model was trained and evaluated utilizing a 4-fold cross validation scheme. Per cross validation split, we used 3 volumes for training and the remaining volume for validation. As quantitative performance metric, the mean squared error between the maximum intensity projection (MIP) OCTA image based on a volume with 8 repetitions per cluster and the same volume, where only a subset of 2 repetitions per cluster was used, was calculated. We compared the mean MSE based on the denoised 2 repetition volume over the 4 cross validation splits with the MSE of the non-denoised 2-repetition volume and the 8-repetition volume. Results show an improvement of the mean MSE from 61.7 based on the non-denoised 2-repetition volumes to a mean MSE of 16.6 based on the denoised 2-repetition volumes.

To denoise a widefield volume it is first flattened at the IS/OS (described in 2B) and cropped from the first non-zero pixel to the IS/OS. Afterwards, patches of $32 \times 32 \times 32$ voxels are inferred one after the other. Due to the nature of a convolution and the filter size of $3 \times 3 \times 3$ the voxels on the border of a patch are not used, resulting in an effective patch-size of $30 \times 30 \times 30$. Therefore, the pointer is moved 30 voxels after denoising one patch. On the used GPU a batch-size of 64 was chosen, being able to infer one line in x at once. After inference, the patches of $30 \times 30 \times 30$ are copied into the output-volume. No further assembling steps are performed. Depending on the thickness of the retina, the size of a whole volume is approximately $150 \times 2048 \times 2048$ voxels, leading to $\sim$25,000 patches per volume.

After denoising the en face projections are created with a maximum intensity projection over the whole depth.

As a quality metric the contrast-to-noise ratio (CNR) is used, which was calculated as follows:

$$CNR = \frac{f - b}{\sqrt{\delta_f^2 + \delta_{fb}^2}} \qquad (6)$$

where f and b are the mean gray values for foreground and background and δ_f and δ_b are the standard deviations. As the background a dark non-perfused area is chosen, i.e., the foveal avascular zone. For the foreground a well-perfused healthy area with high capillary density is chosen.

D. Montage Algorithm

The montaging of two OCTA volumes to generate an ultra-widefield angiogram is based on feature detection and matching [45]. En face registration (x, y) between the two volumes is estimated by feature points detected independently in their corresponding retinal angiograms by finding Shi and Tomaso corners [46]. Each point is assigned to a scale-invariant feature transform descriptor before the algorithm searches for matching feature pairs. The best transformation between the two images is determined using a random sample consensus (RANSAC) algorithm. Axial registration (z) is computed by examining layer locations in the two volumes. The en face and axial resulting registrations are mapped to an enlarged FOV encompassing the region covered by the two OCTA volumes, extrapolating registrations in non-overlapping regions considering distortion effects.

III. RESULTS

A. Denoising

Figure 5 shows a 9 mm $\times$ 4.5 mm acquisition of a diabetic patient acquired. The volume was acquired using $n = 8$ repetitions in approximately 5 seconds of acquisition time. To evaluate the performance of the denoising algorithm the volume was processed four times, using only the first two repetitions along with denoising, using four and all eight without denoising. A standard processing procedure would include Gaussian filtering to suppress noise. Therefore, a three-dimensional Gaussian filter was applied to the volumes, except for the denoised one. After performing the denoising algorithm no further filtering is needed. For the en face projections the CNR was calculated as described above.

Even though the implemented denoising algorithm has not been optimized towards runtime efficiency yet, denoising of a single volume takes approximately one minute on a single GPU (Nvidia 1060 GTX).

With the presented algorithm a similar CNR using only 2 repetitions was obtained compared to using 4 repetitions without denoising. Here several effects might play a role. First, the network was trained with even higher quality data, i.e., data using 8 repetitions instead of 4. Further, the algorithm smoothens vascular structure heavily, having a positive effect on visual appearance.

We believe that the three-dimensional nature of the network contributes significantly to the obtained excellent image quality. Operating on the three-dimensional data, the network
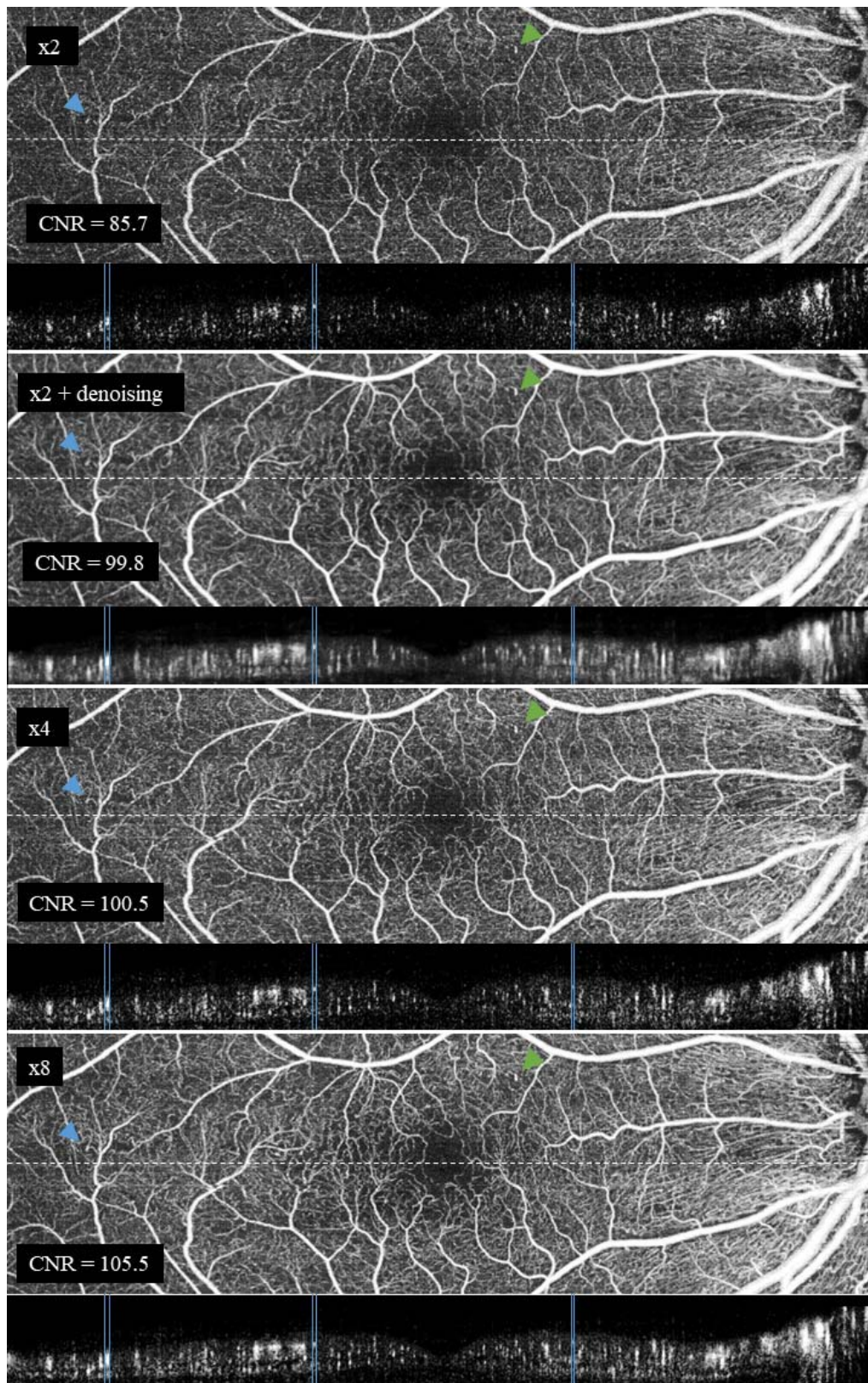
Fig. 5. Comparison of OCTA images of the same volume of a patient with mild diabetic retinopathy acquired with the presented prototype, covering a FOV of 9 mm × 4.5 mm, acquired with 8 repetitions per slow axis position in approximately 5 seconds and processed using different steps. The first two rows are processed only the first two repetitions and the denoising algorithm was applied to the one in the second. Third and fourth row was processed with either 4 or all 8 repetitions. To create the en face projections maximum intensity was used. The contrast to noise ratio was calculated using the dark area of the foveal avascular zone and an area with high vascular density on the right side. The blue and green triangles indicate microaneurysms. To proof that the resolution and vascular features can be preserved when applying the denoising algorithm the blue lines surround certain vessels and have the same positions at all B-scans.

considers a significantly larger amount of information to distinguish vascular structure from background compared to performing image enhancement on just 2D enface slices.

A major concern of applying AI image enhancement algorithms to medical data sets is of course the possibility of the network "inventing" features or in our case capillaries.

We therefore acquired data sets with 4 repetitions and again reconstructed OCTA en face images using all 4 repetitions but without denoising and using the same data set reconstructed

OCTA en face images using only the first two repetitions plus the presented denoising. We did not find a single case of the network hallucinating vessels, nor any instances of the network removing vessels that were visible in the ground truth images. Because the algorithm works on small patches, made-up vessels would appear very prominent due to sharp borders from one patch to the next one. After imaging over 100 patients, we did still not come across such an artifact.

The resolution seems to be maintained after denoising, at least compared to OCTA images processed with a higher number of repetitions. This is shown in Figure 5, which shows a slice through a part of the retina containing many vessels with a diameter in the magnitude of the optical resolution. Although, a slight broadening can be noticed after denoising, this is may however be due to the speckle pattern still present in the raw 2 repetition images. When comparing the denoised image to one created from a higher number of repetitions no change in capillary diameter is observable. Also, small microaneurysms indicated by the blue and green arrows, which represent some of the smallest clinically relevant features in our images, are well maintained after denoising.

The reduction in B-scan repetitions from 4 to two halves the acquisition time from at least 15 to at least 8 seconds. This improves patient comfort as well as image quality tremendously. Especially when imaging patients with deprived ability to fixate or poor general compliance it is best to keep the acquisition time short despite active retinal tracking. Retinal tracking can correct for lateral motion, by rescanning areas where motion occurred, but it leads to an exponential increase in acquisition time as patients tend to have more difficulty fixating with increasing duration of the acquisition. Eventually, patients get so nervous that they move their head around leading to large axial motion that cannot be easily corrected with our current system or accommodate to closer distances resulting in defocused images with reduced SNR. The reduction of the acquisition time to at minimum approx. 8 seconds (depending on patient's motion) was therefore a key enabling factor to consistently obtain high quality ultra-widefield scans across a broad patient population.

B. Ultra-Widefield OCTA

The optical lateral resolution of commercially available OCT devices is around 15 to 25 μm (1/e^2). To achieve Nyquist sampling for a 12 mm $\times$ 12 mm scan pattern, approximately 2 million A-scans are needed. With a common A-scan rate of 100 kHz this would lead to a minimum acquisition time of more than 20 seconds. In order to limit the acquisition time, widefield scans on today's commercial instruments are
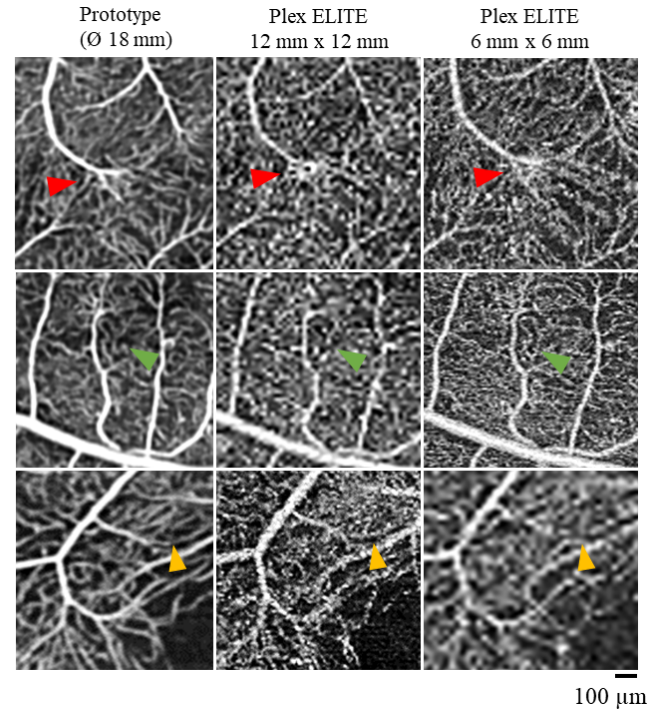

Fig. 6. Comparison of OCTA en face projections of different acquisitions of diabetic patients with a FOV of 1 mm x 1 mm acquired with the presented prototype (2 repetitions + denoising) and two separate acquisitions done with a commercial device.

TABLE III
SAMPLING OF DIFFERENT IMAGE

DEVICE	FOV [mm]	Points	Spacing [μm]
Prototype	Ø 18	2048 x 2048	8.7
Plex ELITE	12 x 12	800 x 800	15
	6 x 6	600 x 600	10

therefore typically under sampled with respect to the Nyquist-Shannon sampling theorem. Because some features, such as microaneurysms, are of the size of the optical resolution, an under sampled volume may lead to reduced detectability of such features. Therefore, typically at least two scans are acquired in clinical practice: a widefield scan to provide an overview, followed by one or more smaller, denser sampled scans of regions of interest. This results in a longer overall procedure time with the need to realign the patient between scans. Due to our extremely high A-scan rate, we were able to maintain Nyquist sampling even for widefield scans with a diameter of 18 mm while keeping the acquisition time at a reasonable minimum of 8 seconds per eye. Figure 6 shows a comparison between magnified regions of interest (approx. 1 mm $\times$ 1 mm) from our widefield prototype and smaller scans from a state-of-the-art commercial SS-OCTA device of the same eye. A comparison between the different sampling specifications is given in Table III.
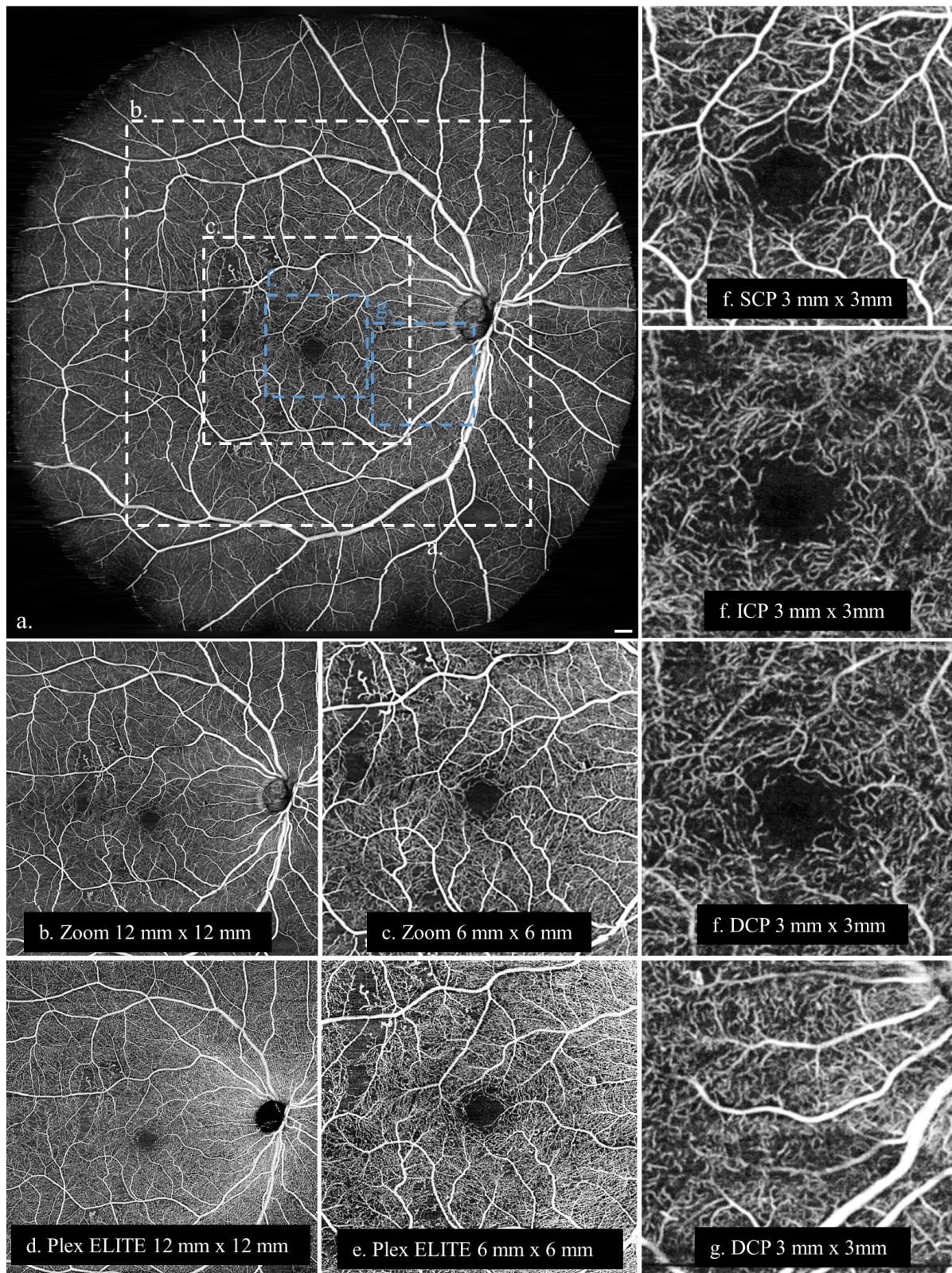
Fig. 7. OCTA en face projection of a diabetic patient with a FOV of 18 mm in diameter (a) and two Zooms (b), (c) with a FOV of 12 mm × 12 mm and 6 mm × 6 mm. For comparison (d) and (e): two OCTA en face projections of separate volumes taken with a commercially available device and corresponding FOVs; The right column shows the different plexi (superficial capillary plexus (SCP), intermediate capillary plexus (ICP) and deep capillary plexus (DCP)) from an area around the fovea (f) and the DCP from an area next to the ONH (g). Scale bars are 1 mm.

Due to our dense sampling many features are better resolved in the images acquired with the widefield prototype despite its much larger FOV. In the first row the red arrow indicates a vessel branch. It is well resolved in the widefield image acquired with the prototype. But in the images taken with the commercial device, especially in the 12 mm × 12 mm
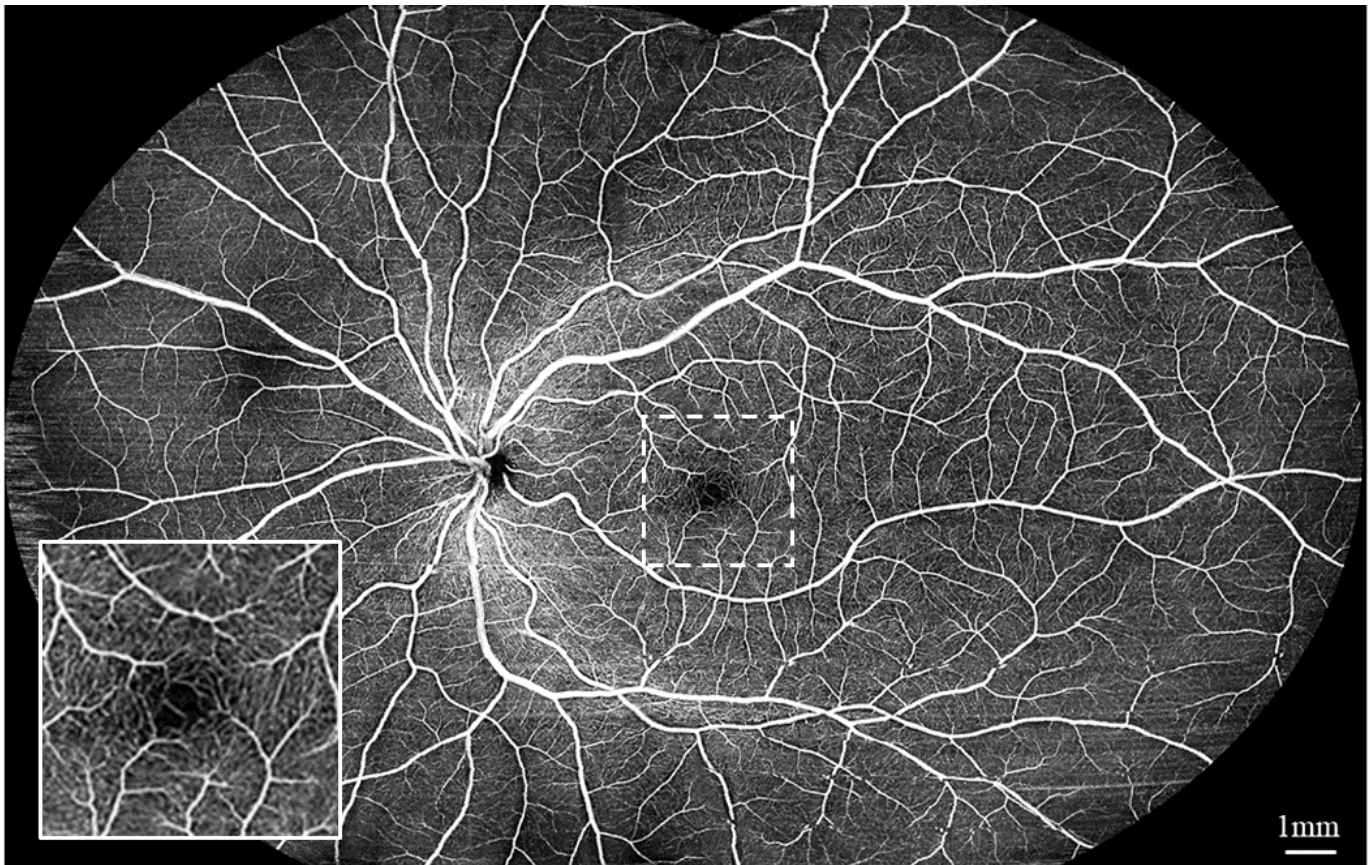
Fig. 8. Dual shot widefield OCTA enface projection (maximum intensity) of a healthy subject, where two acquisitions (Ø 18 mm) are stitched together to cover a FOV of approximately 23 mm × 17 mm on the retina (top and bottom slightly cropped). The overall procedure time is below one minute. On the left bottom corner is a zoom of the foveal zone with a FOV of 3 mm × 3 mm.

it cannot be resolved and appears as an undefined white cluster. This can probably be attributed to the lower sampling resolution in combination with a post processing step, like for example a Gaussian filter. On a larger zoom level, it could be misinterpreted as a microaneurysm and therefore contribute to a misdiagnosis.

In the other two rows small capillaries are indicated by the green and yellow arrows. Again, they are well resolved as individual capillaries by the presented widefield prototype.

However, due to insufficient sampling of the 12 mm × 12 mm scan, these capillaries appear as a white cluttered area, despite identical optical resolution of the two systems. In regions with low signal such structures may be very difficult to distinguish from noise. In Figure 7 a representative clinical case of a diabetic patient is presented. The acquisition time when imaging patients was slightly longer, due to eye motion leading to more frequent rescans than in healthy subjects, but never exceeded 30 seconds. The data recording for the included images in particular took less than 20 seconds, depending on patient's motion, which was typical. Because these patients received a comprehensive eye exam as part of the regular clinical routine prior to being imaged with the widefield OCTA prototype, their pupils were dilated using two drops of tropicamid 1% (Mydriaticum Agepha).

The widefield OCTA projection of the retina is seen in Fig. 7a. Two magnified views of the areas indicated in (a) are shown in (b) and (c). Two separate acquisitions of the same eye taken with the commercial device with a corresponding native FOV of 12 mm × 12 mm and 6 mm × 6 mm are shown in (d) and (e).

Further, two views of the same volume segmented with an experimental multilayer segmentation algorithm, still in development, [47] are shown in the right column. Since we stored the raw data of all volumes, we will be able to separate the different layer, once the development of the algorithm is completed.

The first three images (f) show the area around the fovea split into the three layers: superficial capillary plexus (SCP) from inner limiting membrane (ILM) to inner plexiform layer (IPL), intermediate capillary plexus (ICP) from IPL to the inner nucleus layer (INL) and the deep capillary plexus (DCP) from INL to the outer plexiform layer (OPL). The second view shows the DCP from an area next to the ONH (g).

The minimum acquisition time for a 12 mm × 12 mm scan in the fastest configuration (200 kHz) on the commercial device is approximately 8 seconds, being the same as for the larger and much denser sampled FOV of the shown prototype. As already shown in Figure 6, small features are better resolved with the herein presented widefield prototype.

Even compared to the denser sampled 6 mm × 6 mm acquisition, the prototype images are visually more pleasing due to the smoother, more continuous appearance of the capillaries after application of the denoising step. With the presented prototype only one acquisition per eye was sufficient to perform a wide-ranging examination on the capillary level. With the large FOV vessel dropouts in the periphery could be detected, which would have been missed using today's standard FOVs.

By acquiring two widefield OCTA volumes of 18 mm in diameter with offset fixation points and montaging them we extended our FOV further to 23 mm × 18 mm (Figure 8). This allows us to also examine areas in the far periphery. The sub-acquisitions are taken with the same settings as previously described, except that the fixation target for the patient were offset by +15 degree and −15 degree horizontally. The presented montaged ultra-widefield OCTA image in Figure 8 is therefore still Nyquist sampled, providing capillary level resolution over the entire FOV, which can be seen at the magnified view with a FOV of 3 mm × 3mm of the foveal region placed in the bottom left corner. The shadows in the bottom part of the image were caused by the patient's eye lashes, which is unfortunately a common artifact in ultra-widefield fundus images and OCT scans.

Even when having to realign the patient between sub-acquisitions the overall procedure time can typically be kept well under two minutes. As mentioned in the introduction, state-of-the-art commercial instruments already offer similar montaged scans, but due to their cumbersome acquisition process, they are hardly used in clinical practice. With the shortened exam time of the presented prototype, capturing ultra-widefield scans becomes for the first time feasible in a standard clinical setting.

IV. Conclusion

We developed a clinical ophthalmic SS-OCTA system operating at MHz A-scan rates, i.e., 8 to 16 times faster than state of the art commercial systems. It allowed us to acquire very densely sampled widefield OCTA volumes in a reasonable amount of time. This enabled for the first time the resolution of the finest capillaries all the way out to the periphery of a 23 mm field of view. With the presented OCTA denoising algorithm, we were able to compensate for the lower SNR caused by the extreme acquisition speed and even exceed the angiography image quality of much slower state-of-the-art commercial systems. The presented AI network is the first to operate on volumetric OCT data.

We have already successfully imaged >100 diabetic retinopathy patients with the presented instrument as part of a clinical study. Next steps include a further expansion of the single shot FOV and the addition of further processing steps of the OCTA image reconstruction pipeline.

References

[1] S. B. Kim, E. J. Lee, J. C. Han, and C. Kee, "Comparison of peripapillary vessel density between preperimetric and perimetric glaucoma evaluated by OCT-angiography," *PLoS ONE*, vol. 12, no. 8, Aug. 2017, Art. no. e0184297, doi: 10.1371/journal.pone.0184297.

[2] R. B. Rosen et al., "Earliest evidence of preclinical diabetic retinopathy revealed using optical coherence tomography angiography perfused capillary density," *Amer. J. Ophthalmol.*, vol. 203, pp. 103–115, Jul. 2019, doi: 10.1016/j.ajo.2019.01.012.

[3] G.-S. Ying et al., "Association of baseline characteristics and early vision response with 2-year vision outcomes in the comparison of AMD treatments trials (CATT)," *Ophthalmology*, vol. 122, no. 12, pp. 2523–2531, Dec. 2015, doi: 10.1016/j.ophtha.2015.08.015.

[4] N. Quinn et al., "The clinical relevance of visualising the peripheral retina," *Prog. Retinal Eye Res.*, vol. 68, pp. 83–109, Jan. 2019, doi: 10.1016/j.preteyeres.2018.10.001.

[5] W. S. Lim, G. Grimaldi, L. Nicholson, K. Basheer, and R. Rajendram, "Widefield imaging with clarus fundus camera VS slit lamp fundus examination in assessing patients referred from the national health service diabetic retinopathy screening programme," *Eye*, vol. 35, no. 1, pp. 299–306, Jan. 2021, doi: 10.1038/s41433-020-01218-x.

[6] A. Nagiel, R. A. Lalane, S. R. Sadda, and S. D. Schwartz, "Ultra-widefield fundus imaging," *Retina*, vol. 36, no. 4, pp. 660–678, Apr. 2016, doi: 10.1097/IAE.0000000000000937.

[7] C. E. Pang, V. P. Shah, D. Sarraf, and K. B. Freund, "Ultra-widefield imaging with autofluorescence and indocyanine green angiography in central serous chorioretinopathy," *Amer. J. Ophthalmol.*, vol. 158, no. 2, pp. 362–371, Aug. 2014, doi: 10.1016/j.ajo.2014.04.021.

[8] S. Patel, A. Shi, T. Wibbelsman, and M. Klufas, "Ultra-widefield retinal imaging: An update on recent advances," *Therapeutic Adv. Ophthalmol.*, vol. 12, pp. 1–12, Jan. 2020, doi: 10.1177/2515841419899495.

[9] C. Blatter et al., "Ultrahigh-speed non-invasive widefield angiography," *J. Biomed. Opt.*, vol. 17, Jan. 2012, Art. no. 070505, doi: 10.1117/1.JBO.17.7.070505.

[10] L. Ginner, C. Blatter, T. Schmoll, M. Gröschl, and R. Leitgeb, "Widefield OCT angiography at 400 KHz utilizing spectral splitting," *Photonics*, vol. 1, pp. 369–379, Jan. 2014, doi: 10.3390/photonics1040369.

[11] R. Poddar and J. S. Werner, "Implementations of three OCT angiography (OCTA) methods with 1.7 MHz a-scan rate OCT system on imaging of human retinal and choroidal vasculature," *Opt. Laser Technol.*, vol. 102, pp. 130–139, Jun. 2018, doi: 10.1016/j.optlastec.2017.12.033.

[12] M. Salas et al., "Visualization of micro-capillaries using optical coherence tomography angiography with and without adaptive optics," *Biomed. Opt. Exp.*, vol. 8, p. 207, Dec. 2016, doi: 10.1364/BOE.8.000207.

[13] J. Xu, S. Song, Y. Li, and R. K. Wang, "Complex-based OCT angiography algorithm recovers microvascular information better than amplitude- or phase-based algorithms in phase-stable systems," *Phys. Med. Biol.*, vol. 63, no. 1, Dec. 2017, Art. no. 015023, doi: 10.1088/1361-6560/aa94bc.

[14] S. Makita, F. Jaillon, M. Yamanari, M. Masahiro, and Y. Yasuno, "Comprehensive in vivo micro-vascular imaging of the human eye by dual-beam-scan Doppler optical coherence angiography," *Opt. Exp.*, vol. 19, pp. 83–1271, Jan. 2011, doi: 10.1364/OE.19.001271.

[15] W. Drexler, U. Morgner, R. Ghanta, F. Kärtner, J. Schuman, and J. Fujimoto, "Ultrahigh-resolution ophthalmic optical coherence tomography," *Nature Med.*, vol. 7, pp. 502–507, May 2001, doi: 10.1038/86589.

[16] A. Fercher, "Optical coherence tomography-development, principles, applications," *Zeitschrift für Medizinische Physik*, vol. 20, pp. 76–251, Nov. 2010, doi: 10.1016/j.zemedi.2009.11.002.

[17] A. Fercher, C. Hitzenberger, W. Drexler, G. Kamp, and H. Sattmann, "In vivo optical coherence tomography," *Amer. J. Ophthalmol.*, vol. 116, pp. 113–114, Jan. 1993, doi: 10.1016/S0002-9394(14)71762-3.

[18] C. Hitzenberger, "Measurement of the axial eye length by laser Doppler interferometry," *Investigative Ophthalmol. Vis. Sci.*, vol. 32, pp. 616–624, Jan. 1991.

[19] C. Hitzenberger, P. Trost, P.-W. Lo, and Q. Zhou, "Three-dimensional imaging of the human retina by high-speed optical coherence tomography," *Opt. Exp.*, vol. 11, pp. 2753–2761, Jan. 2003, doi: 10.1364/OE.11.002753.

[20] E. A. Swanson et al., "In vivo retinal imaging by optical coherence tomography," *Opt. Lett.*, vol. 18, pp. 1864–1866, Jan. 1993, doi: 10.1364/OL.18.001864.

[21] R. A. Leitgeb, R. M. Werkmeister, C. Blatter, and L. Schmetterer, "Doppler optical coherence tomography," *Prog. Retinal Eye Res.*, vol. 41, pp. 26–43, Jul. 2014, doi: 10.1016/j.preteyeres.2014.03.004.

[22] T. Schmoll, I. R. Ivascu, A. S. G. Singh, C. Blatter, and R. A. Leitgeb, "Intra- and inter-frame differential Doppler optical coherence tomography," *Sovremennye tehnologii V Med.*, vol. 7, no. 1, pp. 34–43, Mar. 2015, doi: 10.17691/stm2015.7.1.05.

[23] T. Schmoll, C. Kolbitsch, and R. Leitgeb, "Ultra-high-speed volumetric tomography of human retinal blood flow," *Opt. Exp.*, vol. 17, pp. 4166–4176, Apr. 2009, doi: 10.1364/OE.17.004166.

[24] B. Baumann et al., "Total retinal blood flow measurement with ultrahigh speed swept source/Fourier domain OCT," *Biomed. Opt. Exp.*, vol. 2, pp. 1539–1552, Jun. 2011, doi: 10.1364/BOE.2.001539.

[25] J. Fingler, D. Schwartz, C. Yang, and S. Fraser, "Mobility and transverse flow visualization using phase variance contrast with spectral domain optical coherence tomography," *Opt. Exp.*, vol. 15, pp. 12636–12653, Nov. 2007, doi: 10.1364/OE.15.012636.

[26] R. Wang et al., "Wide-field optical coherence tomography angiography enabled by two repeated measurements of B-scans," *Opt. Lett.*, vol. 41, p. 2330, May 2016, doi: 10.1364/OL.41.002330.

[27] Carl Zeiss Meditec. *PLEX Elite 9000 From ZEISS*. Accessed: Dec. 2, 2020. [Online]. Available: https://www.zeiss.com/meditec/int/product-portfolio/optical-coherence-tomography-devices/plex-elite-9000-swept-source-oct.html

[28] R. Leitgeb, C. K. Hitzenberger, and A. F. Fercher, "Performance of Fourier domain vs. time domain optical coherence tomography," *Opt. Exp.*, vol. 11, no. 8, pp. 889–894, Apr. 2003, doi: 10.1364/OE.11.000889.

[29] R. Huber, M. Wojtkowski, and J. G. Fujimoto, "Fourier domain mode locking (FDML): A new laser operating regime and applications for optical coherence tomography," *Opt. Exp.*, vol. 14, no. 8, pp. 3225–3237, 2006, doi: 10.1364/OE.14.003225.

[30] W. Wieser, B. Biedermann, T. Klein, C. Eigenwillig, and R. Huber, "Multi quality 3D imaging at 20 million a-scans and 4.5 GVoxels per second," *Opt. Exp.*, vol. 18, pp. 14685–14704, Jul. 2010, doi: 10.1364/OE.18.014685.

[31] T. Klein, W. Wieser, L. Reznicek, A. Neubauer, A. Kampik, and R. Huber, "Multi-MHz retinal OCT," *Biomed. Opt. Exp.*, vol. 4, pp. 1890–1908, Oct. 2013, doi: 10.1364/BOE.4.001890.

[32] J. P. Kolb, T. Klein, C. L. Kufner, W. Wieser, A. S. Neubauer, and R. Huber, "Ultra-widefield retinal MHz-OCT imaging with up to 100 degrees viewing angle," *Biomed. Opt. Exp.*, vol. 6, no. 5, p. 1534, May 2015, doi: 10.1364/BOE.6.001534.

[33] X. Wei, T. T. Hormel, Y. Guo, T. S. Hwang, and Y. Jia, "High-resolution wide-field OCT angiography with a self-navigation method to correct microsaccades and blinks," *Biomed. Opt. Express*, vol. 11, pp. 3234–3245, 2020. [Online]. Available: https://opg.optica.org/boe/fulltext.cfm?uri=boe-11-6-3234&id=431970

[34] X. Wei, T. T. Hormel, Y. Guo, and Y. Jia, "75-degree non-mydriatic single-volume optical coherence tomographic angiography," *Biomed. Opt. Exp.*, vol. 10, no. 12, p. 6286, Dec. 2019, doi: 10.1364/BOE.10.006286.

[35] S. Kadomoto, A. Uji, Y. Muraoka, T. Akagi, and A. Tsujikawa, "Enhanced visualization of retinal microvasculature in optical coherence tomography angiography imaging via deep learning," *J. Clin. Med.*, vol. 9, no. 5, p. 1322, May 2020, doi: 10.3390/jcm9051322.

[36] L. Xi et al., "A deep learning based pipeline for optical coherence tomography angiography," *J. Biophotonics*, vol. 12, pp. 1–10, Jun. 2019, doi: 10.1002/jbio.201900008.

[37] Z. Jiang et al., "Comparative study of deep learning models for optical coherence tomography angiography," *Biomed. Opt. Exp.*, vol. 11, pp. 1580–1597, Feb. 2020, doi: 10.1364/BOE.387807.

[38] *Safety of Laser Products—Part 1: Equipment Classification and Requirements*, document IEC 60825-1:2007, 2007.

[39] *Safety of Laser Products—Part 1: Equipment Classification and Requirements*, document IEC 60825-1:2014, 2014.

[40] *Ophthalmic Instruments—Fundamental Requirements and Test Methods—Part 2: Light Hazard Protection*, Standard ISO 15004-2:2007, 2007.

[41] S. Makita, T. Fabritius, and Y. Yasuno, "Full-range, high-speed, high-resolution 1 micron spectral-domain optical coherence tomography using BM-scan for volumetric imaging of the human posterior eye," *Opt. Exp.*, vol. 16, pp. 20–8406, Jul. 2008, doi: 10.1364/OE.16.008406.

[42] T. Schlegl et al., "Data-centric AI approach to improve optic nerve head segmentation and localization in OCT en face images," 2022, *arxiv:2208.03868*.

[43] S. Chiu, X. Li, P. Nicholas, C. Toth, J. Izatt, and S. Farsiu, "Automatic segmentation of seven retinal layers in SDOCT images congruent with expert manual segmentation," *Opt. Exp.*, vol. 18, pp. 28–19413, Aug. 2010, doi: 10.1364/OE.18.019413.

[44] W. J. Choi et al., "Characterizing relationship between optical microangiography signals and capillary flow using microfluidic channels," *Biomed. Opt. Exp.*, vol. 7, no. 7, pp. 2709–2728, Jul. 2016, doi: 10.1364/BOE.7.002709.

[45] L. De Sisternes et al., "OCTA volume combination for montaging and averaging," *Investigative Ophthalmol. Vis. Sci.*, vol. 61, no. 9, 2020, Art. no. PB0086.

[46] J. Shi and C. Tomasi, "Good features to track," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 600, Mar. 2000, pp. 593–600, doi: 10.1109/CVPR.1994.323794.

[47] L. D. Sisternes et al., "Segmentation and characterization of intra-retina layer thickness in widefield SS-OCT: Comparison of normal retina and diabetic retinopathy with/without macular edema," *Investigative Ophthalmol. Vis. Sci.*, vol. 59, no. 9, p. 1512, 2018.