3D Retinal Vessel Segmentation in OCTA Volumes: Annotated Dataset MORE3D and Hybrid U-Net with Flattening Transformation

Julian Kuhlmann[1], Kai Rothaus[2], Xiaoyi Jiang[1], Henrik Faatz[2], Daniel Pauleikhoff[2], and Matthias Gutfleisch[2]

[1] Faculty of Mathematics and Computer Science, University of Münster, Münster, Germany
[2] Department of Ophthalmology, St. Franziskus-Hospital, Münster, Germany

Abstract. Optical Coherence Tomography Angiography (OCTA) extends the 3D structural representation of the retina from conventional OCT with an additional representation of "flow" and is used as noninvasive angiography technique in ophthalmology today. While there are several works for the segmentation of vascular network in OCTA images, most of them are tested on 2D enface images (top view projection) only. Such 2D enface images have the drawback that they depend on a good 3D segmentation of retinal layers, the so-called slabs. Especially in case of retinal diseases (e.g. exudations of the retina) this segmentation is not always clear, even for medical experts. In contrast, we consider the problem of full 3D segmentation of retinal vessels in OCTA images. We present the dataset MORE3D (Münster Octa REtina 3D dataset) that is the first one with 3D annotation. We introduce a general flattening transformation that simplifies and accelerates the 3D data labeling and processing, and also enables a specialized data augmentation. Moreover, we realize a hybrid U-net to achieve a first reference segmentation performance on our dataset. In addition to the common performance metrics we also apply skeleton-based metrics for a more comprehensive structural performance evaluation. With this work we contribute to the advancement of 3D retinal vessel segmentation in OCTA volumes.

Keywords: Optical Coherence Tomography Angiography · 3D annotated dataset · flattening transformation · vascular network segmentation.

1   Introduction

The OCT technology enables deep scanning of the retinal layers and provides a 3D structural representation of the retinal layers. The OCT angiography introduced in 2014 [39] extends the structural OCT by the "flow". For each voxel the variation in time of the reflected laser spectrum is measured [16]. This variation is mainly caused by the travel of erythrocytes in the blood. The resulting OCTA volume represents the "flow". Besides motion artifacts during the image capture, there are projection artifacts and a rather large proportion of white
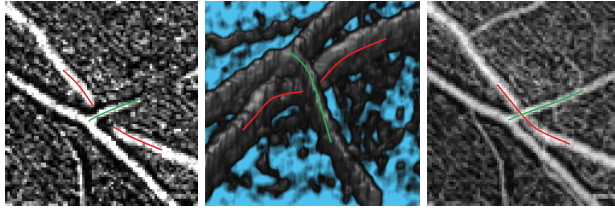
Fig. 1. Visualization of a volume with vessels crossing in different depths. From left to right: single 2D slice, rendered 3D volume, 2D projection (enface image).

noise. The morphology of the vascular network is a suitable biomarker for various pathological changes of the retina [10]. This has been demonstrated especially in quantitative morphological analysis of choroidal neovascularisations (CNV) in neovascular age-related macular degeneration (nAMD), which is the most common cause for legal blindness in the western world.

Currently, most algorithms realize vessel segmentation by a depth-limited 2D projection within anatomical retinal layers. A 2D OCTA enface image is generated pixel-wise by aggregating (e.g. averaging, maximum, or minimum) the flow information of the corresponding voxel stack in the specific retinal layer (slab). Depending on the manufacturer of the OCTA device there are different segmentation layers. In this paper we have used the Avanti™ from OptoVue, which segments the retinal vascular layers into a) the superficial vascular plexus, limited by the inner limiting membrane (ILM) and inner plexiform layer (IPL) -9 μm, and b) the deep vascular plexus, limited by IPL -9 μm and outer plexiform layer (OPL) +9 μm. Furthermore, c) the outer retina and d) the choriocapillaris are defined. For each of these so-called slabs a 2D enface OCTA image is generated by the Avanti™ software. Other OCTA manufactors have slightly different slab definitions but the approach to reduce the data to 2D enface images is the same. This 2D-projection procedure has the drawback that it depends on a good segmentation of morphological retinal layers (slabs). Especially in case of retinal diseases (e.g. exudations of the retina, disordering of retinal layers) the slab segmentation is not always clear, even for medical experts.

In this work we consider full 3D segmentation in OCTA images. Such a segmentation helps distinguish vessels that are merged in 2D images due to projection. For example, vessels in different depths of the retina may be crossing in the enface image (see Figure 1). A full segmentation in 3D OCTA is challenging due to the different sources of noise. Additionally, segmentation inaccuracy affects differently on popular performance metrics for 2D and 3D cases, see the discussion in Section 2.

Related work. There exist a few publicly available datasets with OCTA enface images: ROSE [30], OCTA-500 [21], FAROS [43], see more details in Section 2. However, there is no publicly available OCTA dataset with 3D manual vessel annotations yet.

OCTA dataset	Data	Ground Truth
OCTA-500 [21]	3D volumes	2D pixel level (enface image)
ROSE [30]	3D volumes	2D pixel level (enface image) or centerlines only
FAROS [43]	3D volumes	2D pixel level (enface image)
Our dataset MORE3D	3D volumes	**3D voxel level**

Table 1. Comparison of datasets of OCTA images.

A survey on segmentation and classification methods for OCTA images can be found in [32]. Numerous works exist for vesselness computation [19] and vesselness segmentation in 2D retinal fundus images [18,41]. Recent works have addressed several problems using OCTA, e.g. enhancement [27], quantification of choroidal neovascularization [42], detection of diabetic retinopathy [7], 3D shape modeling of retinal microvasculature [46], and 3D retinal vessel density mapping [35]. Relatively few work deals with the segmentation of vascular network in OCTA images. They verify on 2D enface images only [3,6,8,13,14,22,24,26,28,31] or segment 3D vessels without related ground truth [33,46].

Contributions. Our key contributions are: 1) A flattening transformation dedicated to 3D OCTA images to simplify and accelerate the 3D data labeling and processing. It also enables a specialized data augmentation. 2) A novel 3D OCTA dataset with 3D labeled vessel network. To our knowledge it is the first one of this kind in the literature. 3) A hybrid U-net that achieves a first reference segmentation performance on our dataset based on an integration of vesselness measures and different data sources. 4) In addition to the popular voxel-based performance metrics we also apply skeleton-based evaluation metrics to study structural quality of segmentation. Generally, with our work we also resolve the problem of the dependence on an appropriate pre-segmentation for 2D enface images.

2   2D vs. 3D vessel detection

Here we briefly discuss differences between 2D and 3D vessel detection.

Datasets. In this paper we present the first 3D OCTA dataset MORE3D with *3D labeled vessel network*. As shown in Figure 2, the 3D ground truth provides a much more detailed view of the structure of the vessels: The overall curvation, the different diameters in depth or even small crossings from one vessel above another (compare Figure 1). To our best knowledge, there are three datasets containing 3D OCTA volumes [21,30,43]. But they only include 2D labeled ground truth (see Table 1 for a comparison).

Segmenting and analyzing the retinal vascular structure in 3D offers many advantages in comparison to processing 2D enface images. The retinal vasculature is morphologically differentiated by their (increasing) distance to the ILM in a superficial, intermediate, and deep plexus [20]. Many pathologies affect the plexus differently [1,38]. In practice, medical experts still use individual enface
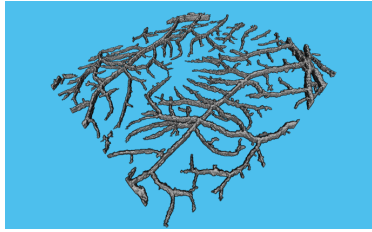
Fig. 2. 3D ground truth.

Fig. 3. Profile of a perfectly round vessel (white) in 2D (left) and 3D (right), where the orange visualizes the area when $1/6$ higher diameter than the real vessel is segmented.

images. With our presented method, the 3D structure of the vessels is made directly available, which simplifies the working of medical experts.

Complexity of vessel detection. Before we introduce the performance metrics in Section 6 we want to emphasize that the absolute values of derived metrics have different significance in 2D (e.g. enface image based processing) and 3D (e.g. our work based on 3D annotation). If we overestimate the diameter r of the vessel by p percent (see Figure 3), then we can determine the True Positive Rate and similarly the Positive Predictive Value if we underestimate the vessel diameter by p percent (see Appendix for the formulas). For example, with $p = 0.2$ we have $TPR_{2D} = 0.8333$ but only $TPR_{3D} = 0.6944$ and $PPV_{2D} = 0.8$ but only $PPV_{3D} = 0.64$ for a vessel like in Figure 3. So if we compare the values of 2D and 3D metrics where TPR or PPV are involved we should take into account that the 3D values are always lower for similar performing algorithm compared to a 2D case. For the False Positive rate (FPR) the influence is much lower, since in retinal vessel segmentation tasks the amount of negatives in the ground truth is much greater than the amount of positives and therefore usually $TN \gg FP$.

3   Flattening transformation for 3D OCTA images

We want to take advantage of the physiology of the eye. The retinal vascular network evolves parallel to the Internal Limiting Membrane (ILM) of the surface of the retina. Thus, we can transform the OCTA image so that the vascular network subsequently evolves in the x/y plane. We call this the "flattening transformation".

Let $I(x, y, z)$ be a 3D OCTA image. We define the 3D vascular network $V(x, y, z)$ and function $f_I(x, y)$ by:

$$V(x, y, z) = \begin{cases} 1, & \text{if vascular network} \\ 0, & \text{otherwise} \end{cases}, \quad f_I(x, y) = \{z, \text{ where } (x, y, z) \in \text{ILM}\}$$

where the latter represents the distance between the top of the OCTA image and the detected ILM, see Figure 4. Then the flattened image I' is defined by:

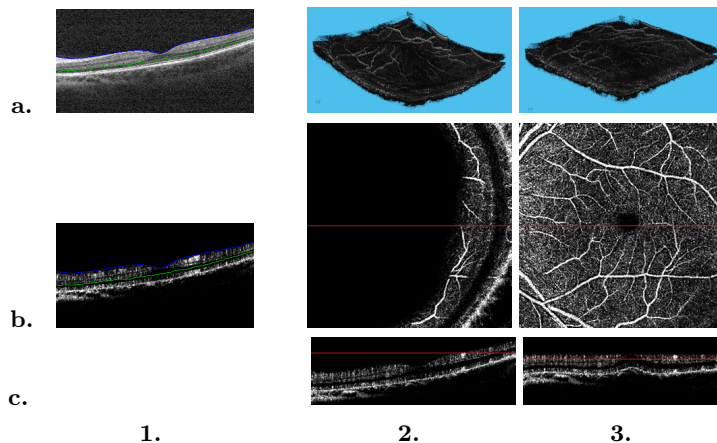$$I'(x, y, z) = I(x, y, z - f_I(x, y) + b) \tag{1}$$

Fig. 4. Flattening transformation. x/z slice of OCT (**a.1.**) and OCTA (**b.1.**) with the detected ILM (blue) and ONL (Outer Nuclear Layer, green). 3D OCTA: original (**a.2.**) and flattened (**a.3.**). Row **b.2./3.**: x/y slice corresponding to the red line in row **c.2./3.**. Row **c.2./3.**: x/z slice corresponding the red line in the row **b.2./3.**.

where $b > 0$ is an offset. The offset leads to a shift so that we still have a part of the vitreous body above the retina new image, see Figure 4. This transformation implies that the ILM in the flattened image is a plane $f_{I'}(x, y) = b = const.$ After the transformation, the vascular network in I' evolves parallel to the x/y plane, see 4. The related inverse transformation of (1) is given by:

$$V(x, y, z) = V'(x, y, z - b + f(x, y))$$

In addition, the flattening transformation considerably simplifies and accelerates the data labeling by a big margin. Without this, the annotator has to label on slices similar to the one in Figure 4.b.2, where the vessels dip in and out of the slices and often the decision if a voxel belongs to the vascular network depends on the slices above or below. With the flattening transformation, instead, the annotator can work on images similar to the one in Figure 4.b.3, which allows much faster decisions and labeling of bigger areas at once.

4   OCTA dataset MORE3D with 3D labeled vessel network

Our dataset consists of 21 OCT and OCTA 3D volumes that were provided by the Department of Ophthalmology, St. Franziskus Hospital, Münster. It will be made available at: https://www.uni-muenster.de/PRIA/forschung/more3d.html. All the OCTA scans were captured with the Avanti™ from Optovue and have a resolution of nearly isotropic 304×304 voxels in the x/y axis and 640 for OCT respectively 160 for OCTA in the z-axis. Hence 4 voxels in the z-axis in OCT
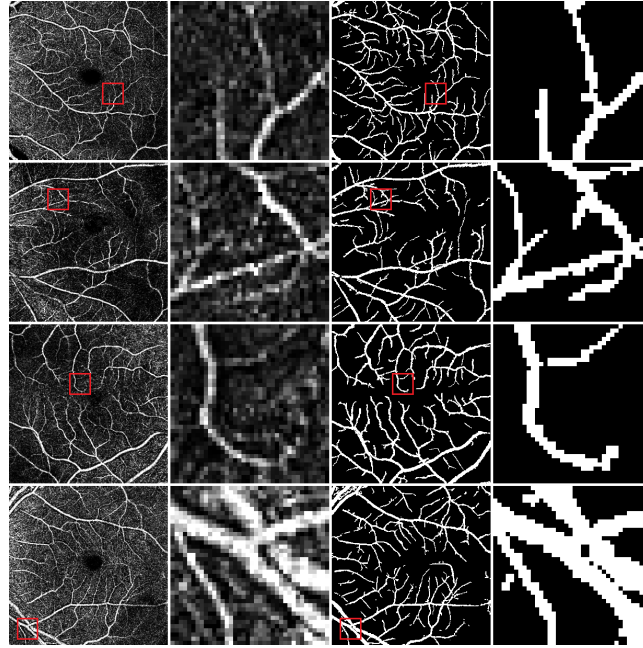
Fig. 5. Illustration of OCTA images (in x/y plane with 1 voxel thickness in z-direction) and their annotations. From left to right: OCTA, enlarged section of the OCTA, Ground Truth, enlarged section of the Ground Truth.

correspond to 1 voxel in OCTA. Since the OCTA images are derived from the OCT image, they are naturally aligned. We found that the vessels are stretched in the z-direction (see [17] for further information). With the definition of the Signal-to-Noise Ratio as $SNR = 10 \cdot \log_{10}(\frac{\mu}{\sigma})$, where μ is the average signal value and σ is the standard deviation, our dataset after normalization to $[0, 1]$ has an $SNR_{\mathrm{OCTA}} = -1.12$. This is substantially lower than the popular 2D DRIVE fundus retinal image data set [40] with $SNR_{\mathrm{Drive}} = 8.07$, thus indicating the much higher complexity of the OCTA segmentation task at hand.

We segmented the ILM and the ONL, whereby the ONL was refined by hand afterwards due to higher difficulty. The Region of Interest (ROI) was set as all voxels in between the detected ILM and ONL. The resulting ROIs contain between about $1.2 - 1.6$ millions voxels depending on the physiology of the eye. This corresponds to an average ROI thickness of 13–17 voxels or ~ 0.16–0.22mm. Each dataset has a labeled vascular network of 42,000–125,000 voxels (average 76,000 voxels), which amounts to 3–10% of the ROI. The average cross-section area for the vessels is $13.86px \pm 8.73px$, which equals around $4px \pm 2px$ as average vessel diameter.

The superficial and deep vascular network was first hand labeled by the first author (9 datasets) or a student assistant (12 datasets). Afterwards it was corrected by a medical expert. These double graded segmentations are taken as

ground truth (GT). The time needed to label one dataset was about 8–12 hours plus 1–3 hours for the corrections by the medical expert. As one can see in Figure 5, the dataset provides various challenges, very bright vessels, bright background noise to very low intensity and small vessels next to noisy background. The hand labeling was performed on the flattened OCTA images in the x/y plane. The proposed flattening transformation considerably increased the labeling speed and reduced the difficulty to recognize the vessels. The average gradient/slope of the ILM is $5.9\% \pm 3.9\%$ with a median of 5.1%. Thus, the flattening transformation will not disrupt the original structure.

The only needed parameter for the flattening transformation is the ILM. Since this transformation is a voxel-wise shift by the position of the ILM, small mistakes in the ILM result in small errors in the transformation, as long as the segmented ILM stays smooth. Furthermore, the ILM is easily detectable in the OCT image so that a smooth ILM for the OCTA image is given due to the lower resolution in the z-axis.

5   3D vascular network segmentation

5.1   Vesselness measures

A common approach to segmenting vessels is the use of vesselness measures. We consider 12 vesselness measures, which can be divided into three groups: Multiscale filter based (Ricci [34], Chaudhuri [4], Läthèn [29], Azzopardi [2]), morphological operation (Zana [45], Sazak [37]), and multiscale eigenvalue based (Jerman [15], Frangi[11], Sato[36], Li [23], Erdt [9], Zhou[47]). A selection of such measures will be integrated into the segmentation network.

5.2   Hybrid U-net architecture

The decision if a voxel belongs to the background or to a vessel is of local nature. Thus, we choose a patch-based CNN approach to reduce the computational cost and increase the speed. We choose an input patch size as $28 \times 28 \times 5$ and our output patch size $16 \times 16 \times 1$ in $x/y/z$ due to the following reasons. We added 6 voxels in both x and y direction for the input patch so that for each voxel in the output patch we know at least the 13×13 local surrounding in the x/y plane. During labeling it was observed that most voxels could be labeled by viewing the current slice in the x/y plane and for the challenging voxels it was enough to scroll through 1-2 extra slices in the z-direction to decide between vessel and background. Thus, we chose the z-size of the input patch as 5 so that we include 2 slices above and 2 below the output patch location. To reduce imbalance between foreground and background voxels we only learn with patches whose $16 \times 16 \times 1$ GT contains at least 1 foreground voxel.

In this work we focus on integrating vesselness measures and different data sources for boosting segmentation performance. We will study which vesselness measures are the best to be used (see Section 6) and at which position of the hybrid U-net we should insert vesselness measures to perform best. Neural networks
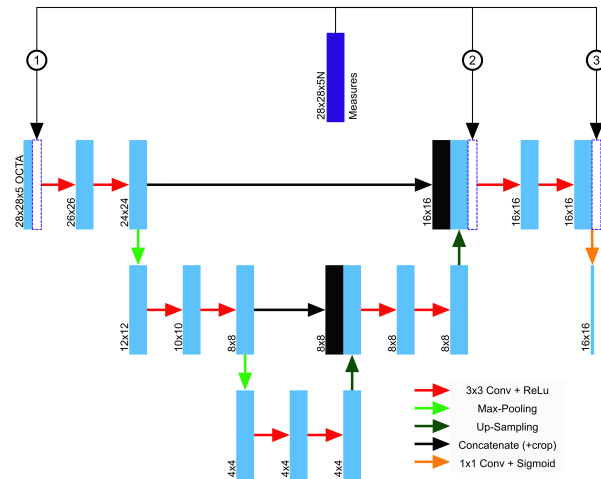
Fig. 6. Hybrid U-net with 3 options to insert vesselness measures: At the front (1), in the middle (2), and at the end (3). The input is 3D of size $28 \times 28 \times (5 \cdot (N+1))$, where N is the number of vesselness measures used as input. After the first 3D convolution we continue with a 2D U-net, starting with 64 channels, double the channels with every max pooling layer, and then reduce to half of the number of channels with every up-sampling. The voxel-wise weighted loss is not shown here.

like U-net certainly have some capacity of learning the vesselness. But providing such information explicitly may simplify the network learning. Note that there are other ways of utilizing vesselness measures. For instance, the Frangi-net [12] reformulates the 2D Frangi vesselness measure into a pre-weighted neural network, which leads to a network that is equivalent to the multi-scale Frangi filter.

Insertion position of vesselness measures. There are potentially 3 positions to insert N vesselness measures into the U-net (see Figure 6; we do not consider multiple insertion positions necessary): a) Front: We concatenate the vesselness measures directly to the OCTA, resulting in an increased input patch size $28 \times 28 \times (5 \cdot (N+1))$. b) Middle: We insert the vesselness measures directly after the upwards path of the U-net, but before the next convolutional layer. The size of the enhanced image patches will be cropped to $16 \times 16 \times (5 \cdot N)$. c) End: We insert the vesselness measures after the last 3x3 convolution layer but before the final 1x1 convolutional layer with the Sigmoid layer. The size of the enhanced image patches will be cropped to $16 \times 16 \times (5 \cdot N)$. We empirically study which one is most beneficial.

Voxel-wise weighted loss using GT labels. To cope with the imbalance between foreground and background and the fact that smaller vessel are significantly harder to segment we will adjust the voxel-wise weight in the loss function. Two approaches will be studied.

The first approach is vessel diameter based. From the labeled GT we can calculate a vessel diameter for each foreground voxel. We propose a weighting

function f for a given voxel x and a diameter function $d(x)$ as:

$$w_{\text{diam}}(x, \alpha, n) = \begin{cases} 1, & \text{if } x \text{ is in background} \\ \alpha \left(\frac{d_{\max}}{d(x)} \right)^n, & \text{otherwise} \end{cases}$$

where $d_{\max}$ is the maximum expected vessel diameter and α is a tuneable parameter to increase or decrease the influence of the weights. For $\alpha = 1$ and $n = 0$ we will receive a uniform response. A higher α increases the focus on the foreground to cope with the imbalance. The higher we choose $n > 0$, the more we will emphasize the difficult smaller vessels in comparison to the wider vessels.

The second approach is edge-based. We use the GT to find all foreground voxels F whose 26-neighborhood contains at least one background voxel and all background voxels B whose 26-neighborhood contains at least one foreground voxel. Then

$$w_{\text{edge}}(x, \alpha) = \begin{cases} \alpha, & x \in F \text{ or } x \in B \\ 1, & \text{otherwise} \end{cases}$$

We will compare the proposed voxel-wise weighted losses with three other losses that address the imbalance problem: (1) The Dice loss, (2) the weighted loss, where each class is weighted by $\frac{1}{classsize}$, and (3) the focal loss [25].

5.3   Flattening-based data augmentation

3D pixel-to-pixel labels are labor-intensive and difficult to obtain [24]. Thus, we use data augmentation to make maximum use of our 21 hand labeled 3D volumes. We flip the volume along x/z plane and rotate it around the z axis in $90°$ steps, increasing the amount of volumes by the factor 8. All other rotations or flips are not used since they might create vascular networks that are not feasible in the human eye. To generate sufficient feasible augmented date, we proceed as follows: a) Flatten the OCTA volume with the flattening transformation (1); b) Tilt the OCTA volume by adding a gradient in the x and y axis. By using the flattening transformation and restricting the norm of the gradient in the second step we make sure that we do not tear up the vascular network and still have a feasible representation of the retina. With this technique we are able to further increase amount of augmentation by the factor 25.

6   Experimental results

6.1   Evaluation metrics

Voxel-based evaluation metrics. We use the following popular methods derived from TP, FP, TN, FN, TPR, TNR, FPR and PPV of the confusion matrix:

- F1-score (same as Dice): $F1 = 2(PPV \times TPR)/(PPV + TPR)$
- Area Under ROC Curve (AUC)
- Accuracy: $Acc = (TP + TN)/(TP + TN + FP + FN)$

- Kappa score: $(Acc-p_e)/(1-p_e)$ with $p_e = \frac{(TP+FN)(TP+FP)+(TN+FP)(TN+FN)}{(TP+TN+FP+FN)^2}$
- False Discovery Rate: $FDR = FP/(FP + TP)$
- G-mean score: $\sqrt{Sensitivity \times Specificity}$

Due to the imbalance between foreground and background some evaluation metrics deliver very high values, e.g. AUC and Accuracy [5].

Skeleton-based evaluation metrics. We also use the skeleton similarity metrics [44] to compare the skeleton of the vessel network. For each segment i, a curve similarity cs_i and a thickness similarity ts_i are computed segment-wise between the GT and the segmentation. The skeleton similarity for one segment is then given by: $ss_i = \alpha \cdot cs_i + (1 - \alpha) \cdot ts_i$ where $\alpha \in [0,1]$ is tuneable parameter to weight cs_i and ts_i. The overall skeleton similarity SS is computed by summing up the individual ss_i weighted with the corresponding segment length. Based on SS, weighted TP and FN are computed, while FP and TN are defined as usual. Finally, the skeleton-based sensitivity (rSe), specificity (rSp), and accuracy (rAcc) are derived (see [44] for more details).

6.2   Vesselness measures on OCTA images

We normalized the voxel values to $[0, 1]$ and due to the high noise level in 3D OCTA images we applied a mean filter smoothing. We tested different sizes and found that a filter of size 3×3×3 is well suited. The results of vesselness measures are presented in Table 2. We calculated the result for 4 different combinations: 2D methods on original image, 2D methods on flattened images, 3D methods on original image, and 3D methods on flattened images. We present the average F1-scores on 21 volumes of our OCTA dataset with 3D labeled vessel network for 12 vesselness measures. Other performance measures provided similar results.

The best performing methods are the morphological operations proposed by Zana [45]. In 2D with our flattening technique it outperforms methods and combinations which are not based on morphological operation by 3.9% points with an F1-score of 0.7721. Furthermore, the AUC values for the measure by Zana are the best for every tested combinations. Overall, we selected one best performing method from each category for further use: Ricci [34], Zana [45], Frangi [11], all 2D flattened.

6.3   3D vessel segmentation in OCTA images

We performed a 3-fold cross validation. The optimal threshold to binarize the neural network output is determined on the training data based on F1-score and then applied to the validation data. With data augmentation as described above we have a total amount of $14 \times 8 \times 25 = 2800$ training volumes for each cross validation step. Each volume contains ∼2500-3000 patches with at least one foreground voxel. We used the Adam optimizer for training.

Insertion position of vesselness measures. Table 3 indicates that an insertion at the front performs best in terms of all voxel-based metrics. The middle

Category	Method	2D	flattened 2D	3D	flattened 3D
1)	Ricci	0.7390	0.7438	0.5643	0.5601
	Chaudhuri	0.6909	0.6942	0.6455	0.6514
	Läthèn	0.6840	0.6709	0.6788	0.6902
	Azzopardi	0.6686	0.6309	*	*
2)	Zana	**0.7594**	**0.7721**	**0.7184**	**0.7134**
	Sazak	0.7371	0.7131	0.7000	0.7018
3)	Jerman	0.6626	0.6557	0.6932	0.6886
	Frangi	0.7042	0.7123	0.6931	0.6939
	Sato	0.6864	0.6673	0.5778	0.5809
	Li	0.6173	0.5995	0.6462	0.6483
	Erdt	0.6864	0.6673	0.6963	0.6972
	Zhou	0.6849	0.6705	0.6602	0.6643

Table 2. Study of vesselness measures: F1-scores on 21 images for vesselness measures. Method category: 1) multiscale filter based, 2) morphological operation, 3) multiscale eigenvalue based. *3D implementation not available.

measure position	F1-score	AUC	Acc	G-mean	Kappa	FDR	rSe	rSp	rAcc
front	**0.8547**	**0.9933**	**0.9836**	**0.9149**	**0.8460**	**0.1315**	0.8152	**0.9833**	0.9185
middle	0.8389	0.9906	0.9822	0.9001	0.8292	0.1352	**0.8361**	0.9788	**0.9258**
end	0.8444	0.9917	0.9823	0.9104	0.8350	0.1453	0.7960	0.9818	0.9115

Table 3. Insertion position of vesselness measures in hybrid U-net with all inputs.

position even reduced the performance in most vovel-based metrics compared to the vanilla U-net without added vesselness measures (compare Table 4). In contrast, rSe and rAcc increase. At first glance this is counter-intuitive. However, a look at the results reveals an over-segmentation (i.e. extracting more vessel segments), thus affecting negatively the voxel-based metrics but positively the skeleton-based metrics. We will use the front insertion option in the following.

Segmentation performance by combination of input sources. Table 4 shows the results. The use of vesselness measures consistently increases all metrics. Adding one measure, e.g. Ricci [34], increases the performance for all voxel-based measures. The skeleton-based rSe and rAcc receive best values when adding Zana [45] whereas rSp is best using Ricci [34]. Using all input sources (OCTA, OCT and all three vesselness measures) tops all voxel-based metrics and is close to the best performance in the skeleton-based metrics. Thus, we will only consider this all-input combination for the following experiments.

Voxel-wise weighted loss. After analyzing the labeled data we set the maximum expected diameter to $d_{max} = 7$. A pixel-wise weighted loss with $w_{edge}(x, 10)$ (compare Table 5) increased the performance in every voxel-based metric except G-mean. It achieved the overall best skeleton-based rSp=0.9844.

Figure 7 shows an example using the weighted loss $w_{diam}(x, 2, 2)$. It outperforms all other presented methods and combinations in skeleton-based rSe and

CNN input	F1-score	AUC	Acc	G-mean	Kappa	FDR	rSe	rSp	rAcc
OCTA	0.8428	0.9923	0.9824	0.9064	0.8335	0.1400	0.8160	0.9799	0.9172
OCTA+OCT	0.8434	0.9920	0.9823	0.9079	0.8340	0.1410	0.8156	0.9806	0.9167
OCTA+Ricci [34]	0.8502	0.9928	0.9832	0.9111	0.8413	0.1334	0.8051	**0.9839**	0.9159
OCTA+Zana [45]	0.8476	0.9928	0.9829	0.9078	0.8385	0.1326	**0.8252**	0.9801	**0.9206**
OCTA+Frangi [11]	0.8436	0.9923	0.9826	0.9056	0.8343	0.1368	0.8082	0.9805	0.9146
all combined	**0.8547**	**0.9933**	**0.9836**	**0.9149**	**0.8460**	**0.1315**	0.8152	0.9833	0.9185

Table 4. Segmentation performance by combination of input sources.

loss weighting	F1-score	AUC	Acc	G-mean	Kappa	FDR	rSe	rSp	rAcc
uniform	0.8547	0.9933	0.9836	0.9149	0.8460	**0.1315**	0.8152	0.9833	0.9185
$w_{\mathrm{diam}}(x, 2, 2)$	0.8436	0.9930	0.9822	0.9133	0.8342	0.1530	**0.8667**	0.9701	**0.9306**
$w_{\mathrm{edge}}(x, 10)$	**0.8574**	**0.9938**	**0.9838**	0.9175	**0.8488**	0.1316	0.8144	**0.9844**	0.9203
$w_{\mathrm{edge}}(x, 20)$	0.8514	0.9933	0.9827	**0.9263**	0.8422	0.1616	0.8461	0.9767	0.9268
Weighted classes	0.8327	0.9912	0.9812	0.8985	0.8228	0.1445	0.8163	0.9746	0.9153
Dice loss	0.8559	$(0.9456)^{\dagger}$	0.9836	0.9191	0.8472	0.1375	0.8045	0.9819	0.9169
Focal loss [25]	0.8485	0.9931	0.9828	0.9104	0.8394	0.1357	0.8296	0.9821	0.9245

Table 5. Voxel-wise weighted loss using all inputs. † The result was close to binarized, leading to a low sampling rate for the AUC, which causes a low AUC.

rAcc, but has one of the highest false detection rates. Visually, this setup tends to overestimate the vessel diameter in the z-axis and underestimates the diameter in the x/y plane. This is also indicated by the higher amount of small voxel groups in in Figure 7 (upper row, right image) . These groups belong to vessels whose diameter in the z-axis was overestimated. This explains the lower values in the voxel-based metrics. The use of the proposed voxel-wise weighted losses was able to outperform the three compared losses addressing the imbalance problem.

7 Conclusion

In this work we presented several contributions to 3D retinal vessel segmentation in OCTA volumes. The proposed flattening transformation considerably accelerates the 3D data labeling and enables a specialized data augmentation. As another advantage it may simplify the training of neural networks that are not rotation-invariant, thus potentially making them more difficult to work on the raw curved images. On the other hand, the influence of the flattening transformation on the subsequent analysis is not consistent. Figure 2 reveals that dependent of the operator, the unflattened version may be favorable. We have presented the first OCTA dataset MORE3D with 3D labeled vessel network. It will foster the research in the community. A hybrid U-net was realized to achieve a first reference segmentation performance on our dataset. The presented work helps resolve the problem of the dependence on an appropriate pre-segmentation for 2D enface images that are the dominating source of data in the literature.
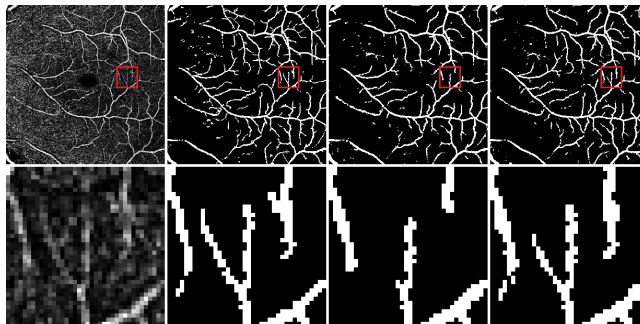
Fig. 7. Illustration of results on a flattened slice. From left to right: OCTA, GT, U-net with OCTA as input, hybrid U-net with all-input combination and voxel-wise weighted loss $w_{\text{diam}}(x, 2, 2)$.

In future we will examine adapting alternate backbone network architectures to full 3D vessel segmentation. Moreover, we will explore transfer learning to efficiently and effectively process OCTA volumes from other manufacturers using the outcome of this work. A fundamental issue is the fact that the absolute values of metrics have different significance for 2D and 3D images (see Section 2). It is thus helpful to find a way that enables to compare the results of 2D and 3D vessel detection problem instances.

Appendix

If we overestimate the diameter r of the vessel by p percent (see Figure 3), then the True Positive Rate will become:

$$TPR_{2\text{D}}(p) = \frac{2r}{2r + 2pr} = \frac{1}{1 + p}$$

$$TPR_{3\text{D}}(p) = \frac{\pi r^2}{\pi r^2 + (\pi(r + pr)^2 - \pi r^2)} = \frac{1}{(1 + p)^2}$$

Similarly if we underestimate the vessel diameter by p percent we have Positive Predictive Value:

$$PPV_{2\text{D}}(p) = \frac{2(r - pr)}{2(r - pr) + 2pr} = 1 - p$$

$$PPV_{3\text{D}}(p) = \frac{\pi(r - pr)^2}{\pi(r - pr)^2 + (\pi r^2 - \pi(r - pr)^2)} = (1 - p)^2$$

These results can be used to compare the inherent complexity of vessel detection in 2D vs. 3D (see Section 2).

Acknowledgments

References

1. Ashraf, M., Sampani, K., Clermont, A., Abu-Qamar, O., Rhee, J., Silva, P.S., Aiello, L.P., Sun, J.K.: Vascular density of deep, intermediate and superficial vascular plexuses are differentially affected by diabetic retinopathy severity. Investigative Ophthalmology & Visual Science **61**(10), 53 (2020)
2. Azzopardi, G. et al.: Trainable COSFIRE filters for vessel delineation with application to retinal images. Medical Image Analysis **19**(1), 46 – 57 (2015)
3. Breger, A. et al.: Blood vessel segmentation in en-face OCTA images: a frequency based method. CoRR **abs/2109.06116** (2021)
4. Chaudhuri, S. et al.: Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans. on Medical Imaging **8**(3), 263–269 (1989)
5. Davis, J., Goadrich, M.: The relationship between precision-recall and roc curves. In: 23rd Int. Conf. on Machine Learning (ICML). p. 233–240 (2006)
6. Eladawi, N. et al.: Automatic blood vessels segmentation based on different retinal maps from OCTA scans. Computers in Biology and Medicine **89**, 150–161 (2017)
7. Eladawi, N. et al.: Early signs detection of diabetic retinopathy using optical coherence tomography angiography scans based on 3d multi-path convolutional neural network. In: IEEE Int. Conf. on Image Processing (ICIP). pp. 1390–1394 (2019)
8. Engberg, A.M.E. et al.: Automated quantification of retinal microvasculature from OCT Angiography using dictionary-based vessel segmentation. In: 23rd Conf. on Medical Image Understanding and Analysis (MIUA). pp. 257–269 (2019)
9. Erdt, M., Raspe, M., Suehling, M.: Automatic hepatic vessel segmentation using graphics hardware. In: Med. Imag. Augment. Reality. pp. 403—412 (2008)
10. Faatz, H. et al.: Optical coherence tomography angiography of types 1 and 2 choroidal neovascularization in age-related macular degeneration during anti-vegf therapy: Evaluation of a new quantitative method. Eye **33**(9), 1466–1471 (2019)
11. Frangi, A.F. et al.: Multiscale vessel enhancement filtering. In: First Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 130–137 (1998)
12. Fu, W., Breininger, K., Schaffert, R., Ravikumar, N., Würfl, T., Fujimoto, J., Moult, E., Maier, A.K.: Frangi-net. In: Workshop Bildverarbeitung für die Medizin. pp. 341–346. Springer Vieweg (2018)
13. Hao, J. et al.: Retinal structure detection in OCTA image via voting-based multi-task learning. IEEE Trans. on Medical Imaging **41**(12), 3969–3980 (2022)
14. Hu, K. et al.: Joint-seg: Treat foveal avascular zone and retinal vessel segmentation in OCTA images as a joint task. IEEE Trans. on Instrumentation and Measurement **71**, 1–13 (2022)
15. Jerman, T. et al.: Enhancement of vascular structures in 3d and 2d angiographic images. IEEE Trans. on Medical Imaging **35**(9), 2107–2118 (2016)
16. Jia, Y. et al.: Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. Optics Express **20**(4), 4710–4725 (2012)
17. Kuhlmann, J., Rothaus, K., Jiang, X., Heimes-Bussmann, B., Faatz, H., Book, M., Pauleikhoff, D.: Axial stretching of vessels in the retinal vascular plexus with 3D OCT-angiography. Translational Vision Science and Technology **11**, 21 (2022)
18. Kumar, K.S., Singh, N.P.: Analysis of retinal blood vessel segmentation techniques: a systematic survey. Multimedia Tools and Applications **82**(5), 7679–7733 (2023)
19. Lamy, J. et al.: Vesselness filters: A survey with benchmarks applied to liver imaging. In: 25th Int. Conf. on Pattern Recognition (ICPR). pp. 3528–3535 (2020)

20. Lavia, C., Bonnin, S., Maule, M., Erginay, A., Tadayoni, R., Gaudric, A.: Vessel density of superficial intermediate and deep capillary plexus using optical coherence tomography angiography. Retina **39**(2), 247–258 (Feb 2019)
21. Li, M., Huang, K., Xu, Q., Yang, J., Zhang, Y., Ji, Z., Xie, K., Yuan, S., Liu, Q., Chen, Q.: OCTA-500: A retinal dataset for optical coherence tomography angiography study. CoRR **abs/2012.07261** (2020)
22. Li, M., Zhang, W., Chen, Q.: Image magnification network for vessel segmentation in OCTA images. In: 5th Chinese Conf. on Pattern Recognition and Computer Vision (PRCV). pp. 426–435. arXiv (2022)
23. Li, Q., Sone, S., Doi, K.: Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans. Medical Physics **30**(8), 2040—-2051 (2003)
24. Li, M. et al.: Image projection network: 3d to 2d image segmentation in OCTA images. IEEE Trans. on Medical Imaging **39**(11), 3343–3354 (2020)
25. Lin, T. et al.: Focal loss for dense object detection. IEEE Trans. on Pattern Analysis and Machine Intelligence **42**(2), 318–327 (2020)
26. Liu, X. et l.: OCTA retinal vessel segmentation based on vessel thickness inconsistency loss. In: IEEE Int. Conf. on Image Processing (ICIP). pp. 2676–2680 (2022)
27. Liu, Y. et al.: Projection artifact suppression for inner retina in OCT angiography. In: IEEE 16th Int. Symposium on Biomedical Imaging (ISBI). pp. 592–596 (2019)
28. Liu, Y. et al.: Disentangled representation learning for OCTA vessel segmentation with limited training data. IEEE Trans. on Medical Imaging **41**(12), 3686–3698 (2022)
29. Läthén, G., Jonasson, J., Borga, M.: Blood vessel segmentation using multi-scale quadrature filtering. Pattern Recognition Letters **31**, 762–767 (2010)
30. Ma, Y. et al.: ROSE: A retinal OCT-Angiography vessel segmentation dataset and new model. IEEE Trans. on Medical Imaging **40**(3), 928–939 (2021)
31. Ma, Z. et al.: Retinal OCTA image segmentation based on global contrastive learning. Sensors **22**(24), 9847 (2022)
32. Meiburger, K.M. et al.: Automatic segmentation and classification methods using Optical Coherence Tomography Angiography (OCTA): A review and handbook. Applied Sciences **11**, 9734 (2021)
33. Pissas, T. et al.: Deep iterative vessel segmentation in OCT Angiography. Biomedical Optics Express **11**(5), 2490 (2020)
34. Ricci, E., Perfetti, R.: Retinal blood vessel segmentation using line operators and support vector classification. IEEE Trans. on Medical Imaging **26**(10), 1357–1365 (2007)
35. Sarabi, M.S. et al.: 3d retinal vessel density mapping with OCT-Angiography. IEEE Journal of Biomedical and Health Informatics **24**(12), 3466–3479 (2020)
36. Sato, Y. et al.: Tissue classification based on 3d local intensity structures for volume rendering. IEEE Trans. on Medical Imaging **6**(2), 160–180 (2000)
37. Sazak, C., Nelson, C.J., Obara, B.: The multiscale bowler-hat transform for blood vessel enhancement in retinal images. Pattern Recognition **88**, 739–750 (2019)
38. Spaide, R.F., Klancnik, J.M., Cooney, M.J.: Retinal vascular layers in macular telangiectasia type 2 imaged by optical coherence tomographic angiography. JAMA Ophthalmology **133**(1), 66–73 (2015)
39. Spaide, R.F. et al.: Optical coherence tomography angiography. Progress in Retinal and Eye Research **64**, 1–55 (2018)
40. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. Medical Imaging **23**(4), 501–509 (2004)

41. Sule, O.O.: A survey of deep learning for retinal blood vessel segmentation methods: Taxonomy, trends, challenges and future directions. IEEE Access **10**, 38202–38236 (2022)
42. Taibouni, K. et al.: Automated quantification of choroidal neovascularization on optical coherence tomography angiography images. Computers in Biology and Medicine **114**, 103450 (2019)
43. Xiao, P., Hu, X., Ma, K., Wang, G., Feng, Z., Huang, Y., Yuan, J.: OMSN and FAROS: OCTA microstructure segmentation network and fully annotated retinal OCTA segmentation dataset. CoRR **abs/2212.13059** (2022)
44. Yan, Z., Yang, X., Cheng, K.: A skeletal similarity metric for quality evaluation of retinal vessel segmentation. IEEE Trans. on Medical Imaging **37**(4), 1045–1057 (2018)
45. Zana, F., Klein, J.: Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. IEEE Trans. on Image Processing **10**(7), 1010–1019 (2001)
46. Zhang, J. et al.: 3d shape modeling and analysis of retinal microvasculature in OCT-Angiography images. IEEE Trans. on Medical Imaging **39**(5), 1335–1346 (2020)
47. Zhou, C. et al.: Automatic multiscale enhancement and segmentation of pulmonary vessels in ct pulmonary angiography images for cad applications. Medical Physics **34**(12), 4567—4577 (2007)