Optimization-based vessel segmentation pipeline for robust quantification of capillary networks in skin with optical coherence tomography angiography

Malte Casper
Hinnerk Schulz-Hildebrandt
Michael Evers
Reginald Birngruber
Dieter Manstein
Gereon Hüttmann

SPIE.

Optimization-based vessel segmentation pipeline for robust quantification of capillary networks in skin with optical coherence tomography angiography

Malte Casper,[a,b,*] **Hinnerk Schulz-Hildebrandt,**[b,c] **Michael Evers,**[a,b] **Reginald Birngruber,**[b,d] **Dieter Manstein,**[b,d] **and Gereon Hüttmann**[b,c]

[a]Massachusetts General Hospital, Harvard Medical School, Cutaneous Biology Research Center, Department of Dermatology, Charlestown, Massachusetts, United States
[b]Universität zu Lübeck, Institute of Biomedical Optics, Lübeck, Germany
[c]LungenClinic Grosshansdorf, Airway Research Center North, Member of the German Center for Lung Research, Grosshansdorf, Germany
[d]Medical Laser Center Lübeck, Lübeck, Germany

Abstract. Optical coherence tomography angiography (OCTA) provides *in-vivo* images of microvascular perfusion in high resolution. For its application to basic and clinical research, an automatic and robust quantification of the capillary architecture is mandatory. Only this makes it possible to reliably analyze large amounts of image data, to establish biomarkers, and to monitor disease developments. However, due to its optical properties, OCTA images of skin often suffer from a poor signal-to-noise ratio and contain imaging artifacts. Previous work on automatic vessel segmentation in OCTA mostly focuses on retinal and cerebral vasculature. Its applicability to skin and, furthermore, its robustness against imaging artifacts had not been systematically evaluated. We propose a segmentation method that improves the quality of vascular quantification in OCTA images even if corrupted by imaging artifacts. Both the combination of image processing methods and the choice of their parameters are systematically optimized to match the manual labeling of an expert for OCTA images of skin. The efficacy of this optimization-based vessel segmentation is further demonstrated on sample images as well as by a reduced error of derived quantitative vascular network characteristics. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JBO.24.4.046005]

1 Introduction

Capillaries are embedded within organs of the human body. As blood vessels of the smallest caliber, they are an important supply channel for blood, nutrients, gas exchange, and waste disposal of the cells.[1] Changes in architecture of the capillary network can be symptomatic for organ aging, inflammatory processes, disease development, and wound healing.[2]

Using optical coherence tomography angiography (OCTA), three-dimensional (3-D) maps of capillary perfusion are obtained *in vivo* with high resolution. Different methods have been developed for OCTA image reconstruction.[3–9] They are based on the movement of red blood cells within the blood stream. These particles cause a variation in the OCT signal intensity and a shift in frequency of the backscattered light due to the Doppler effect.[6] By contrasting the perfusion with static tissue, a 3-D angiographic image is generated. Often, such an angiographic volume is projected to a two-dimensional (2-D) image to condense the angiographic image and to avoid the appearance of image artifacts, such as projection shadows.[10] Such an *en face* maximum-intensity projection (MIP) displays the capillary network of a tissue slab at a glance. Color coding can be used to additionally visualize the depth of the vessels.

For the comparison of the capillary structures along a cohort, such as in clinical studies or the preclinical research environment, it is necessary to extract the quantitative characteristics of the vascular network. Quantitative metrics are, for example, the number of vessels, their diameter, and shape, as well as the density and complexity of the vascular network.[5,11] To derive these metrics from an angiographic image, first, a segmentation is required, which classifies each pixel to represent either a vessel or the background. To obtain such a segmentation map, manual labeling by an expert is the gold standard.[5,12–14] Since this is a time-consuming procedure and the outcome might vary from expert to expert, efforts have been focused on the automatic segmentation of the vascular structures.

In the field of computer vision, segmentation is an important task. Often a pipeline composing different image-processing steps is employed. Such a pipeline usually consists of a preprocessing step, which aims to remove noise and to enhance the contrast between foreground and background, a binarization step that separates the object and background, and a post-processing step that refines the segmented object boundaries.[5,11,15]

In computer vision, it is common to evaluate the results of automatic segmentations with a manual ground-truth segmentation of experienced rater.[16–18] The congruence of the ground

---

*Address all correspondence to Malte Casper, E-mail: MJCasper@mgh.harvard.edu

truth and automatic segmentation can be described by the amount of correctly and falsely classified pixels.[5,16–18] Further, to derive quantitative characteristics of the vascular architecture, such as the number of vessels, their curvature, or the complexity of the network, it is common to reduce the binary vascular map to its skeleton,[5] which substitutes each vessel with its 1-pixel thick centerline.

Most of the published algorithms for vessel segmentation and analysis combine these steps of image processing in various ways. For an interested user it is hard to compare the quality of the proposed algorithms and to estimate the segmentation error that a specific method carries. It is also impossible to directly compare findings of different studies based on OCTA with automatic evaluation and to critically assess published results. Furthermore, in dermatology, OCTA imaging is progressively applied by clinicians in office or bedside situations.[19,20] In these real-world applications, imaging artifacts may be introduced due to movement of the bulk tissue, defocus[21] and shading by air bubbles or lesions. Unfortunately, methods proposed in literature to segment and analyze vascular networks have not been evaluated for robustness against such artifacts that are common for OCTA imaging of the skin.[5,11,15,22,23] However, with OCT as a valuable tool for clinicians and scientists, methods for an automated and robust vessel detection are demanded for a reliable quantification.

In our case, OCTA images were obtained of *in-vivo* mouse skin in order to observe the wound healing processes of capillaries after fractional photothermolysis.[24] With this treatment, the vascular architecture changes dramatically and requires a segmentation method that does handle very different appearances and quality of the OCTA images.

This work demonstrates the creation of an optimization-based segmentation pipeline for angiographic images of skin. It is capable of coping with imaging artifacts and especially is suitable for dermatological applications. A variety of state-of-the-art image-processing methods for denoising, contrast enhancement, binarization, and refinement were systematically tested and combined to a pipeline to optimize the quality of the segmentation and the validity of quantitative metrics. Thanks to our optimization-based vessel segmentation (OBVS), we are now able to analyze even low-quality data of a longitudinal mouse study, which would have been discarded otherwise. Although this optimization was done for OCTA images of cutaneous microvasculature, it might be applicable to retinal and cerebral OCTA images.

2 Material and Methods

A subset of 10 different OCTA volumes acquired during a study for cutaneous photoaging in mice was chosen to find the optimal pipeline for vessel segmentation (see Fig. S2 in the Supplementary Material). This subset of OCTA volumes contains examples of different imaging situations that occurred during this study, such as artifact-free images under optimal condition; slightly defocused images due to epidermal thickening; and images corrupted by motion artifacts, lesions, or airbubbles between skin and objective. The sample volumes were chosen to be representative of skin studies in rodents.

To find the optimal segmentation approach (see Sec. 3.2) the set was split into a training group and a test group with five images each.

2.1 Image Acquisition

OCTA imaging was performed using the commercially available spectral-domain OCT scanner TELESTO II (Thorlabs Inc., Newton, New Jersey). The device operates at a central wavelength of 1300 nm and an axial resolution of 4.2 μm in tissue. Using a lens with a lateral resolution of 13 μm (LSM03, Thorlabs Inc.), we acquired two times oversampled images with a voxel size of $6.5 \times 6.5 \times 4.2$ μm^3. Field of view (FOV) was $6 \times 6 \times 4.2$ mm^3 (length $\times$ width $\times$ depth); A-Scan rate was 76 kHz. For further analysis with reasonable cost for manual labeling, the angiographic volumes were cropped to $400 \times 400 \times 50$ px, showing the capillary layer in between $\sim$80 and 300 μm depth in skin. After angiographic image processing (see Sec. 2.2), the volumes were converted to 2-D images as MIPs.

OCTA images were acquired in the dorsal region of nude mice (Foxn1nu). The tissue was immobilized by using a z-spacer (IMM3, Thorlabs Inc.) in front of the scanner with glycerol as immersion fluid. During an imaging session, the mice were anesthetized with isoflurane (3% induction, 1% to 1.5% maintenance) in 80% air and 20% oxygen. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of the Massachusetts General Hospital (Protocol No.: 2015N000170).

2.2 Angiographic Processing of the OCT Data

Angiographic images were generated using the speckle variance method (svOCT), as proposed by Mariampillai et al.,[3,4] using the minimal amount of 2$\times$ repeated B-Scans at one location.

We evaluated numerous other angiographic algorithms and settings, such as optical microangiography (OMAG),[5,6] correlation mapping (cmOCT),[7] complex differential variance (cdvOCT),[8] and complex correlation (ccOCT).[9] Example images of these algorithms are given in Fig. S1 in the Supplementary Material. The contrast and signal-to-noise ratio (SNR) were compared among these methods using formulas such as proposed by Zhang et al.[6] and Lozzi et al.[25] However, as the results of these two assessments contradicted each other, we were unable to clearly identify the best performing method by contrast or SNR. Visually, in our case, the svOCT method generated more detailed maps of capillary perfusion than cmOCT, cdvOCT, and ccOCT, while it appeared less noisy and richer in contrast to OMAG.

All computations, including the optimization of the segmentation described in Sec. 3.2, were performed using MATLAB (R2016b, MathWorks, Natick, Massachusetts) running on a Dell Precision T1700 with an Intel Xeon(R) CPU with four cores (eight threats) at 3.6 GHz.

2.3 Ground-Truth Segmentation

For each of the 10 images, four manual segmentations are created by the same rater using the itkSnap tool.[26] Even though the manual segmentation is the gold standard method for precise labeling, it is also prone to inaccuracy and variations due to some ambiguous pixel and the subjective decision of the rater. Using the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm, as proposed by Warfield et al.,[13] the most likely true, common segmentation is computed by iteratively assessing the rater's performance and weighting each rating session in a combined segmentation. STAPLE
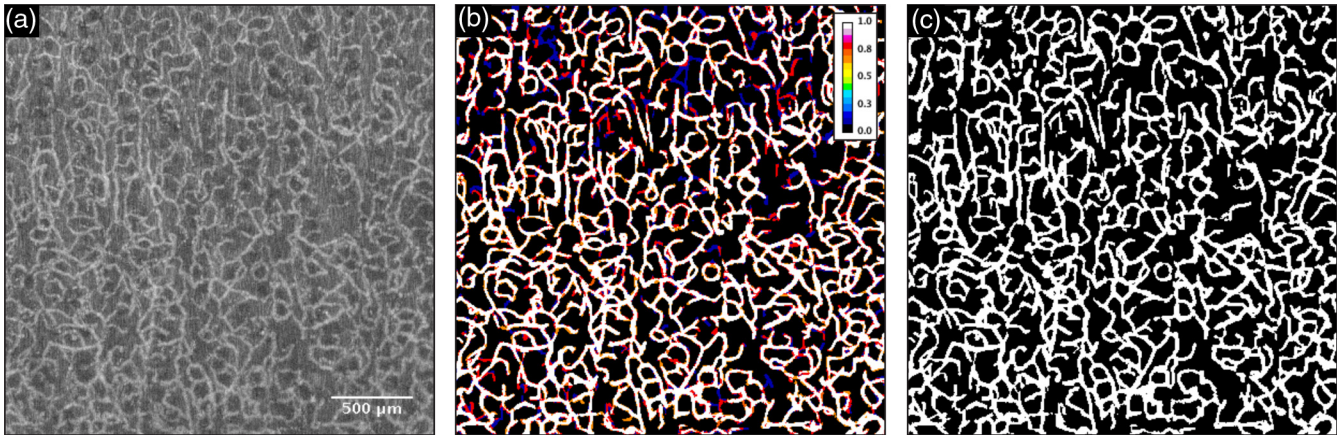
Fig. 1 Illustration of a sample image of size: (a) 400 × 400 px, (b) four manual segmentations are combined using STAPLE and displayed as probability map, and (c) STAPLE result binarized with threshold of 0.75.

terminates with a probability map of every foreground pixel to have been identified correctly.

To obtain a binary mask, we further considered every pixel with probability of 75% and more to be the correct foreground and every other pixel as background (Fig. 1). These segmentations were used as the ground truth to optimize and evaluate the quality of the automatic segmentation algorithms.

2.4 Segmentation Quality Metrics

During the optimization process (see Sec. 3.2), the quality of a segmentation result was evaluated by its congruence with the ground truth.[5,17,27] Each pixel of the binary map was classified as either t_p true-positive, t_n true-negative, f_p false-positive, or f_n false-negative. Based on the ratio of the amount of pixel in each set, the following quality metrics are derived for a segmentation:

$$\text{sensitivity: } Se = \frac{t_p}{t_p + f_n}, \tag{1}$$

$$\text{specificity: } Sp = \frac{t_n}{t_n + f_p}. \tag{2}$$

Youden's index[27]:

$$J = \frac{t_p t_n - f_p f_n}{(t_p + f_n)(f_p + t_n)}, \tag{3}$$

$$= \frac{t_p}{t_p + f_n} + \frac{t_n}{f_p + t_n} - 1, \tag{4}$$

$$= Se + Sp - 1. \tag{5}$$

2.5 Vascular Quantitative Metrics

Quantitative descriptive metrics of vascular networks were obtained from the skeleton of segmentation map to characterize its architecture and condense the information in single values. This is especially helpful when comparing vascular network characteristics among many angiographic images, such as in large datasets of clinical studies. We utilized vascular quantitative metrics as proposed in previous work of OCTA.[5,6,15,17]

The vessel length density (VLD) is given by the amount of skeleton pixel divided by the image size and indicates the number of vessels present in the image.[5,6]

The fractal dimension (FD) quantifies the complexity of a vascular network[28] and is obtainable from both the binary segmentation and the skeletonized map. It has been observed that it gives more characteristic results for skeletons.[5]

To assess the tortuosity of the skeleton, Bullitt et al.[29] proposed to integrate the angular changes along the vessel segments and normalize by their length. This sum of angles metric (SOAM) is effective to recognize abnormalities characterized by high-frequency, "low-amplitude" coils.

To assess these metrics, we implemented the code in MATLAB (R2016b, MathWorks, Natick, Massachusetts). Their provided values were further used to estimate the quality of OBVS and other established methods, as listed in Table 1. The methods were benchmarked against the ground-truth skeletons, which were here obtained by morphological skeletonization[18] of the STAPLE result.

3 Results

3.1 Optimizing Parameters of Four Previously Used Algorithms for Segmentation of Capillaries

Thus far, previous work in OCTA is mainly applied to images of the retinal[11,15,33] and cerebral vasculature. Both show vessel in low scattering environments. The imaging and segmentation of capillaries in skin, as shown in mouse ear by Reif et al.,[5] and in human skin by Liew et al.[22] and Carter et al.,[23] are more challenging, as skin is a stronger scattering imaging environment and is more likely to cause image artifacts.

As a benchmark for OBVS, we have implemented four previously used algorithms for vessel analysis from OCTA images and optimized the respective free parameters for best overall congruence of the five training images with the respective ground-truth segmentation. Table 1 outlines these algorithms together with the parameters, which are chosen by maximizing the Youden's index J for the training datasets in a similar fashion as described for OBVS in Sec. 3.2.

Table 1 Different vessel segmentation algorithm for evaluation of OCTA images, whose performances are compared to OBVS. Methods have been derived from the publications shown. We have optimized the parameters for maximal Youden's index J with respect to the corresponding ground truth for the five training images.

	Methods and parameters for segmentation	Site	Authors
1.	Fixed threshold binarization [optimal: th = 1.066 * mean(OCTA-signal)] skeletonization: Voronoi[30]	Skin	Liew et al.[22]
			Carter et al.[23]
2.	2.1 *k*-means threshold binarization	Retina	Khansari et al.[11]
	2.2 Thickening		
	2.3 Particle filtering (delete objects < 100 px)		
	2.4 Dilation (disk, radius = 1 px)		
	2.5 Filling, skeletonization: distance transform ($r_v = 1$ px)		
3.	3.1 Global threshold for noise removal	Retina	Chu et al.[15]
	3.2 Vesselness[31] ($\sigma_{1,2} = 2$, 10, $\Delta\sigma = 2$, $\beta_{1,2} = 0.5$, 10)		Reif et al.[5]
	3.3 Local adaptive threshold (radius = 9), skeletonization: morphological		
4.	4.1 Niblack binarization[32] (window = 32 px, $\kappa = 0.21$)	Skin	Lozzi et al.[25]
	4.2 Opening (cube, radius = 8), skeletonization: morphological		

3.2 Optimization of Processing Pipeline for OVBS

The segmentation pipeline consisted of the subsequent modules denoising, contrast enhancement, binarization, and refinement. Within each of these stages, several methods were evaluated of their impact on the quality of the result. As indicated in Fig. 2, for each method the input parameter combination was optimized with regard to the Youden's index J, as it combined the specificity with the sensitivity metric. Parameters of rational numbers, such as noise level or filter strength, were optimized using the Nelder–Mead simplex method,[34] whereas integer parameter, such as kernel size, were consecutively set and analyzed. Each optimization step was performed on the whole set of training images.

3.2.1 Denoising

First, to reduce noise and imaging artifacts in the angiographic images, methods of denoising were tested. We evaluated Gaussian filter, median filter, bilateral filter,[35] anisotropic diffusion filter,[36] nonlocal means filter,[37] BM3D filter,[38] and the Hessian-based vesselness filter.[31] Each of these methods was controlled by different input parameter, e.g., kernel size, filter strength, or search area. As indicated in Fig. 2, the optimal combination of input parameter was found using the downhill simplex optimization method[34] and iteration for discrete parameter to maximize the Youden's index J. As we were only able to compare the congruence of two binary maps, the denoised images were provisionally binarized using the optimal global threshold, which was determined for each image individually by another downhill simplex optimization nested into the parameter optimization. Each of these best possible binary maps was then compared to the ground truth to assess the J index.

We observed each of these denoising methods strongly enhancing the quality of the segmentation result compared to

a segmentation without denoising. Especially the BM3D image filter, as proposed by Dabov et al.,[38] scored highest, lifting the average J value from 0.62 without denoising to 0.74 after BM3D filtering. Hence, the BM3D with the optimal noise level parameter $\sigma = 14.6$ was chosen as denoising method for the following investigations.

3.2.2 Contrast enhancement

In the second stage of the segmentation pipeline, after the denoising step with the BM3D filter, methods for contrast enhancement were evaluated. Among methods such as histogram equalization, Retinex,[39] and local-phase-based filtering,[40] the contrast-limited adaptive histogram equalization (CLAHE) led to slightly superior segmentation results. The CLAHE method performed optimally with a number of tiles $n_{\text{tiles}} = 6$ and the contrast enhancement factor $c_{\text{lim}} = 0.6$ with a uniform distribution and was chosen for the following steps.

3.2.3 Binarization

Based on the images that were denoised by the BM3D filter and enhanced by CLAHE, different methods for binarization were evaluated and optimized. The most intuitive approach for binarization is to set a global threshold. This value can either be chosen semiautomatically by the user[15,22,23] or automatically derived from the image by using one of the various known methods.[41] As empirical threshold, such as utilized by Liew and Carter,[22,23] we chose the average of the optimal thresholds for each sample image. These were found with knowledge of the ground-truth segmentation as leading to the maximal Youden's index J. Comparing the results of this fixed threshold with the automatic techniques such as the Otsu's method[42] and Isodata,[43] which adjusted the threshold for each image individually, the Otsu's method led to superior segmentation results. Furthermore, the local adaptive threshold approach, as e.g.,
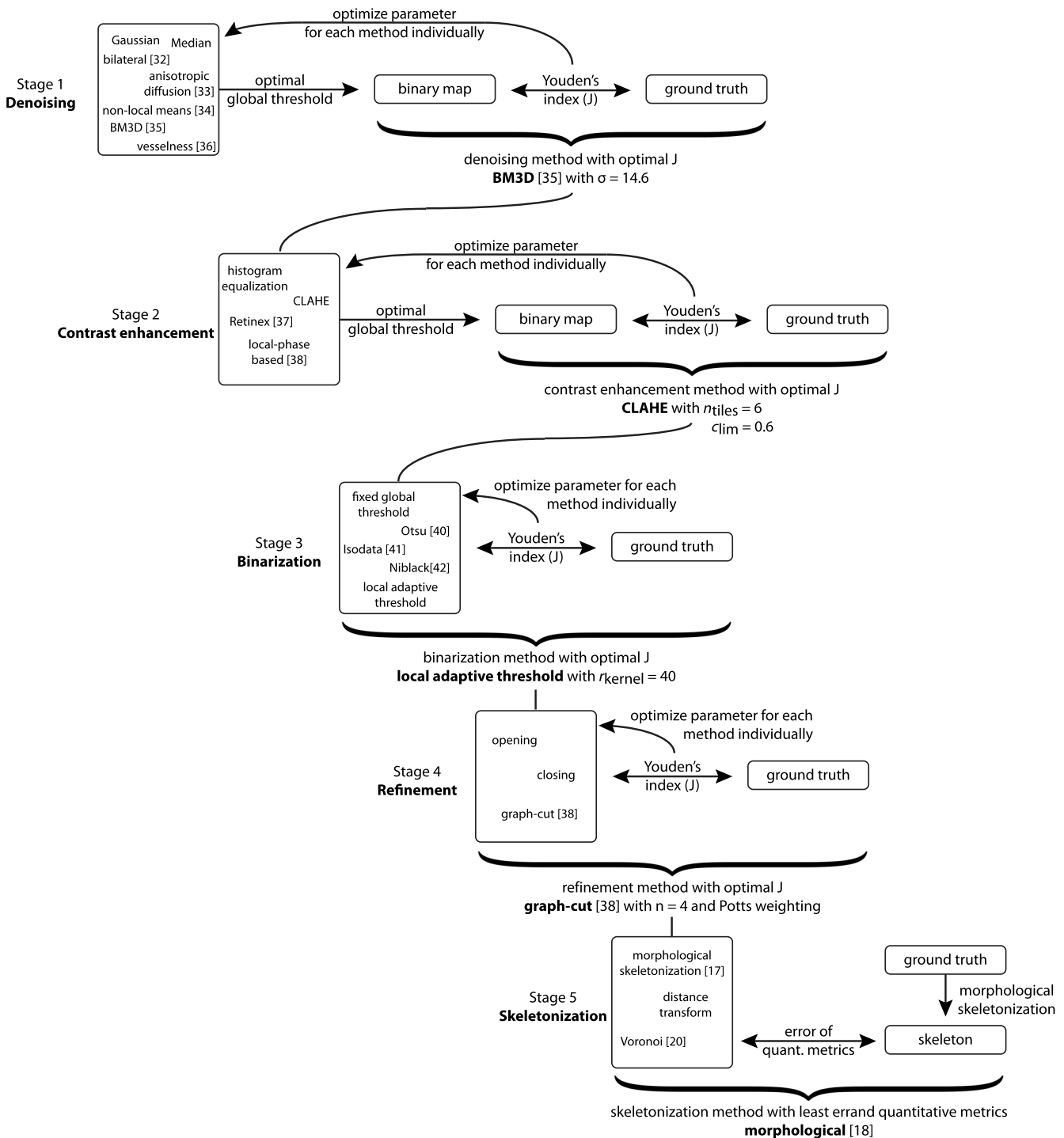
Fig. 2 Scheme for finding the optimal combination of methods and their parameters for capillary segmentation pipeline. In each stage the best-performing method is chosen and taken as base for the following stage. Parameter of each method is optimized using a downhill simplex[34] or iteration, where discrete parameters are required.

utilized by Reif and Chu,[5,15] and the method by Niblack,[32] as proposed for the evaluation of OCTA images by Lozzi et al.,[25] led to better results than the global approaches, when they have been optimized for window size and noise level. Even though for the test dataset the improvement over the global methods was not significant ($p \approx 0.47$), the local adaptive thresholding method with an optimal kernel size of $r = 40$ and a Gaussian statistic with $\sigma = 0.3$ was chosen as the optimal

binarization method, as it was expected to be more robust against artifacts and signal variation.

3.2.4 Refinement

As segmentation in OCTA images was prone to image artifacts, a postprocessing step was conducted to eliminate the small objects of misclassification and to smooth object boundaries

in the binary maps. Here, as methods of refinement, kernel-based morphological operations, such as opening and closing, have been optimized for kernel size and evaluated.

However, in our investigations the improvement of the segmentation result was even more prominent by the min-flow max-cut (graph-cut) approach, such as proposed by Boykov and Kolmogorov.[44,45] This approach solved the active contour model optimization problem, which regularized the length of the object boundary in the binary map in balance with a low intra-object intensity variance of the segmented image. This eventually led to smoother object boundaries and the compensation of minor segmentation errors. Here, the optimization of Youden's index J resulted in a neighborhood size of $n = 4$ weighted by the Potts model with a weighting of foreground and background of $\lambda_f = 1$ and $\lambda_b = 4$, respectively.[17]

3.2.5 Skeletonization

For the analysis of the organization of the segmented vascular network, the binary segmentation map was simplified to its skeleton, which was represented by a line along the center of the vessel. Here, three different methods to obtain such a skeleton were evaluated. The approach proposed by Haralick and Shapiro[18] provided an unregularized skeleton using morphological operations. The resulting skeleton can be further refined by using the distance transform giving the vessel radius at each position on the skeleton.[46] By applying a threshold for the vessel diameter $\varnothing_v$, very small structures were discarded, e.g., here with a diameter of $\varnothing_v < 13\ \mu m$ which is below the lateral imaging resolution of the OCT setup. Moreover, the shape of

a skeleton can be even more regularized using the Voronoi diagram.[30] In our investigations, however, the morphological skeletonization led to optimal results, as the segmentations were smoothened already due to the refinement step and further smoothening of the skeleton was not necessary.

3.3 Comparison of Performance

With the aim to develop a vessel segmentation method robust against artifacts in OCTA imaging of skin, we have benchmarked the proposed OBVS method and the established algorithms listed in Table 1 against the ground truth. In Figs. 3–5, the performance of each of the algorithms is demonstrated for different images of the test group. Furthermore, the segmentation quality metrics of the algorithms applied to the whole group of test images are given in Fig. 6, together with the quantitative metrics such as VLD and FD that are derived from the segmented vascular network and the ground truth.

Figure 3 shows an OCTA image of the test group containing only minor artifacts. A region of 400×400 px was extracted for the evaluation of the segmentation algorithms against the manual ground truth, which was obtained using the STAPLE algorithm (Sec. 2.3). The performance was shown as exemplary at a tile of 100×100 px. Surprisingly, the different methods provided very different results, even though the tile was not corrupted by imaging artifacts and had a contrast common for OCTA images of skin. The binary maps of both the methods, proposed by Liew et al.[22] and Lozzi et al.,[25] based on global and local thresholding, respectively, were affected by noise in the foreground and
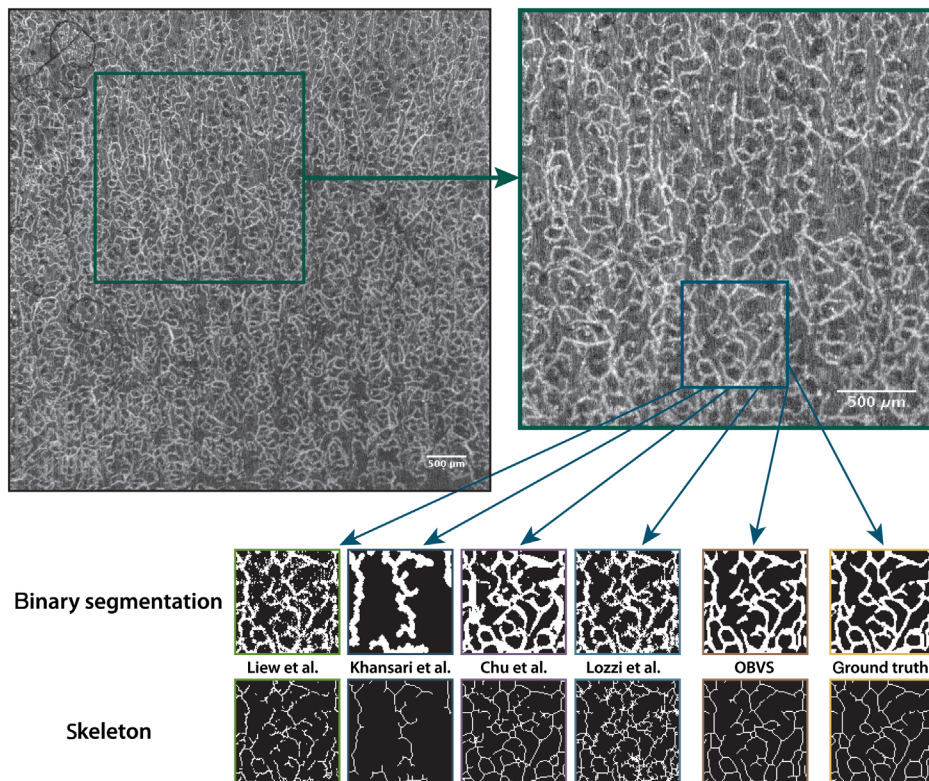
Fig. 3 Performance of the vessel analysis methods on OCTA of mouse skin. The original image is cropped to 400×400 px. For illustration, exemplary results are shown for tiles of 100×100 px (resolution: 6.5 μm). Results using the methods of Liew et al.,[22] Khansari et al.,[11] Chu et al.,[15] Lozzi et al.,[25] and OBVS are shown as segmented vessels as well as skeletons in comparison to the ground truth.
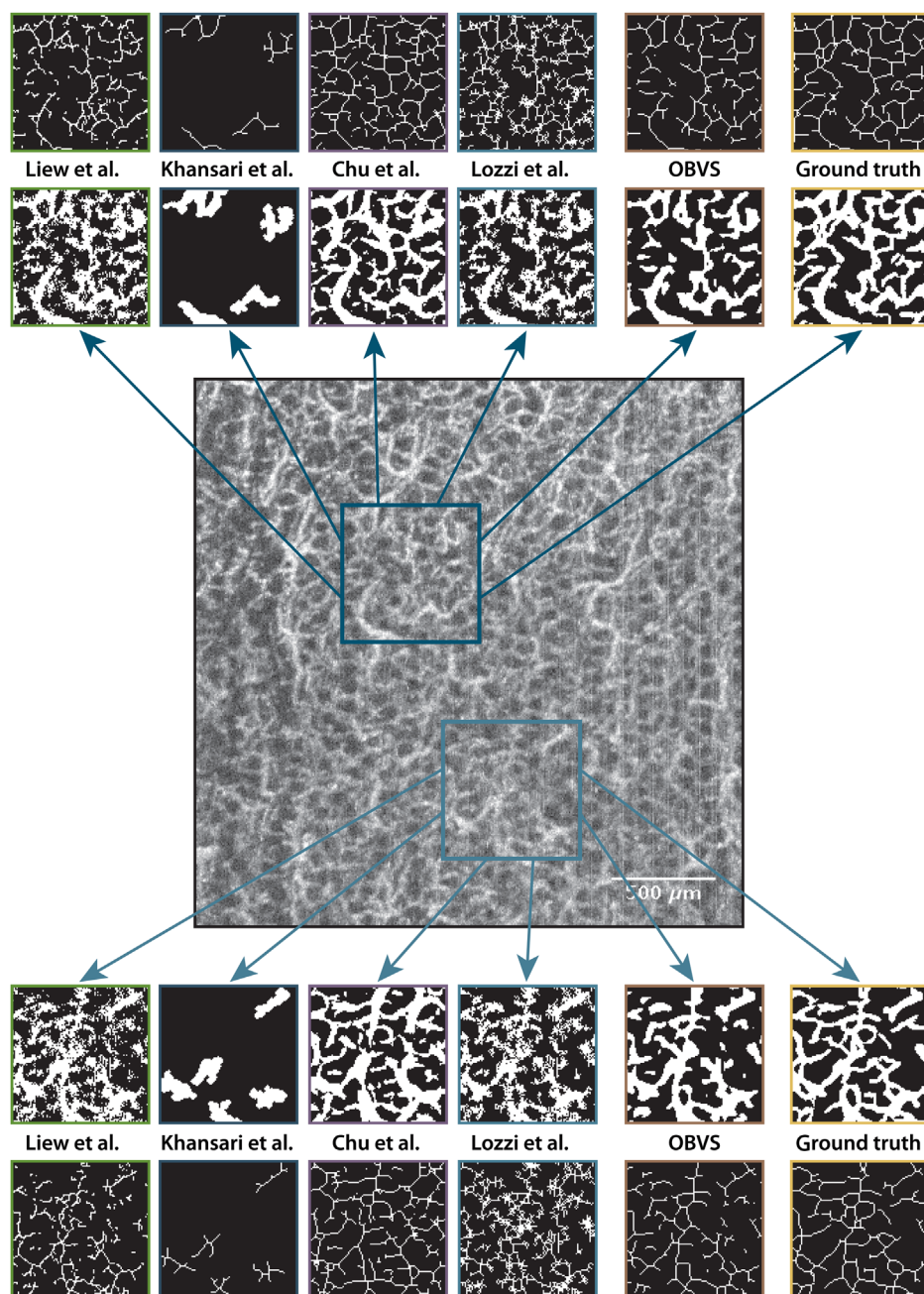
Fig. 4 Performance of vessel segmentation methods on blur in upper and blur and low contrast in lower tile. Slight motion artifacts are visible as vertical lines.

background. The method proposed by Khansari et al.,[11] which mainly consisted of morphological operations, and the method proposed by Chu et al.,[15] applying Hessian-based filtering, appeared to be less sensitive against such noise corruption. However, Khansari's method[11] was developed to detect large vessels and gives just a few vessels with an enlarged caliper. OBVS performed superior compensating the noise successfully, while also maintaining the integrity of vessel diameter, boundary, and connectivity.

The skeleton generated by the method of Liew et al.[22] was obtained using Voronoi regularization[30] and compensated most of the noise successfully, whereas the skeleton method of Chu et al.,[15] which used morphological skeletonization, gave a few vessel fragments in the background. The result of Lozzi et al.[25]

appeared rather toothy and angular. Only OBVS appeared to match the ground truth well, finding the structure and smoothness of the vessel in the ground truth.

This observation could be made as well in Figs. 4 and 5, which contain artifacts. In Fig. 4, OBVS is evaluated on a blurry image that has a poor contrast and is corrupted by sample motion. Figure 5 shows the cutaneous capillary network one day after the fractional photothermolysis treatment with a CO_2 laser.[24] Here, the vascular structures are distributed sparsely with a low contrast in the upper tile and dilated surrounding a lesion in the bottom tile, due to inflammation responses of the skin during wound healing.

The quality of each method in terms of sensitivity, specificity, and Youden's index with the ground-truth segmentation is
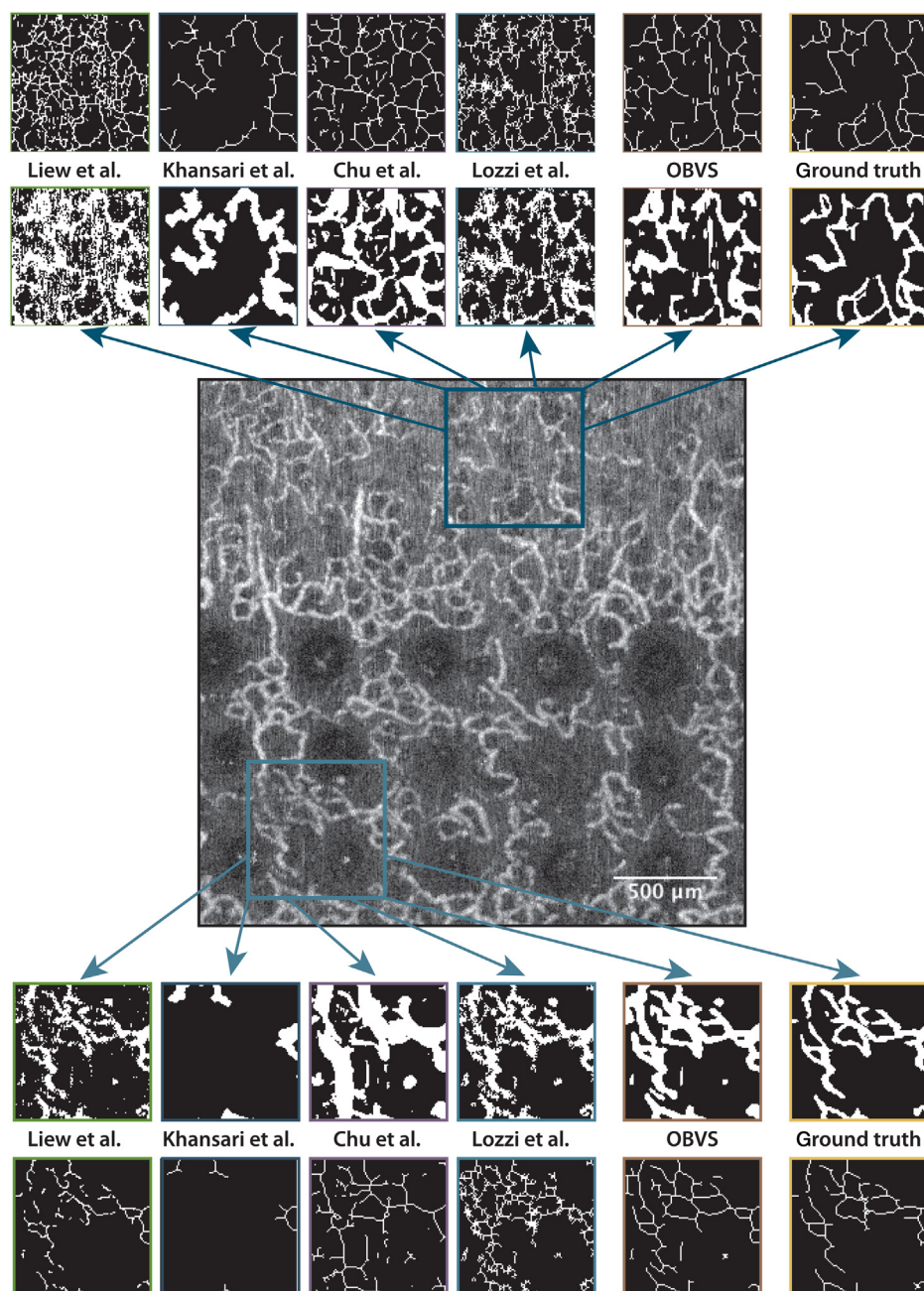
Fig. 5 Performance of vessel segmentation methods on OCTA image of mouse skin capillaries one day after fractional photothermolysis.[24] Imaging is compromised by voids and local inflammation due to wound healing after the laser treatment. Tiles contain sparse vessel structures with low and high contrast, respectively.

averaged over all five test images and is given in Fig. 6(a). The evaluation shown in the plot emphasizes the superiority of OBVS, which scored the highest Youden's index with $J = 0.75 \pm 0.01$. Note, the average congruence of single manual segmentations of the rater compared to the STAPLE result, as described in Sec. 2.3, scored with $J = 0.77 \pm 0.01$.
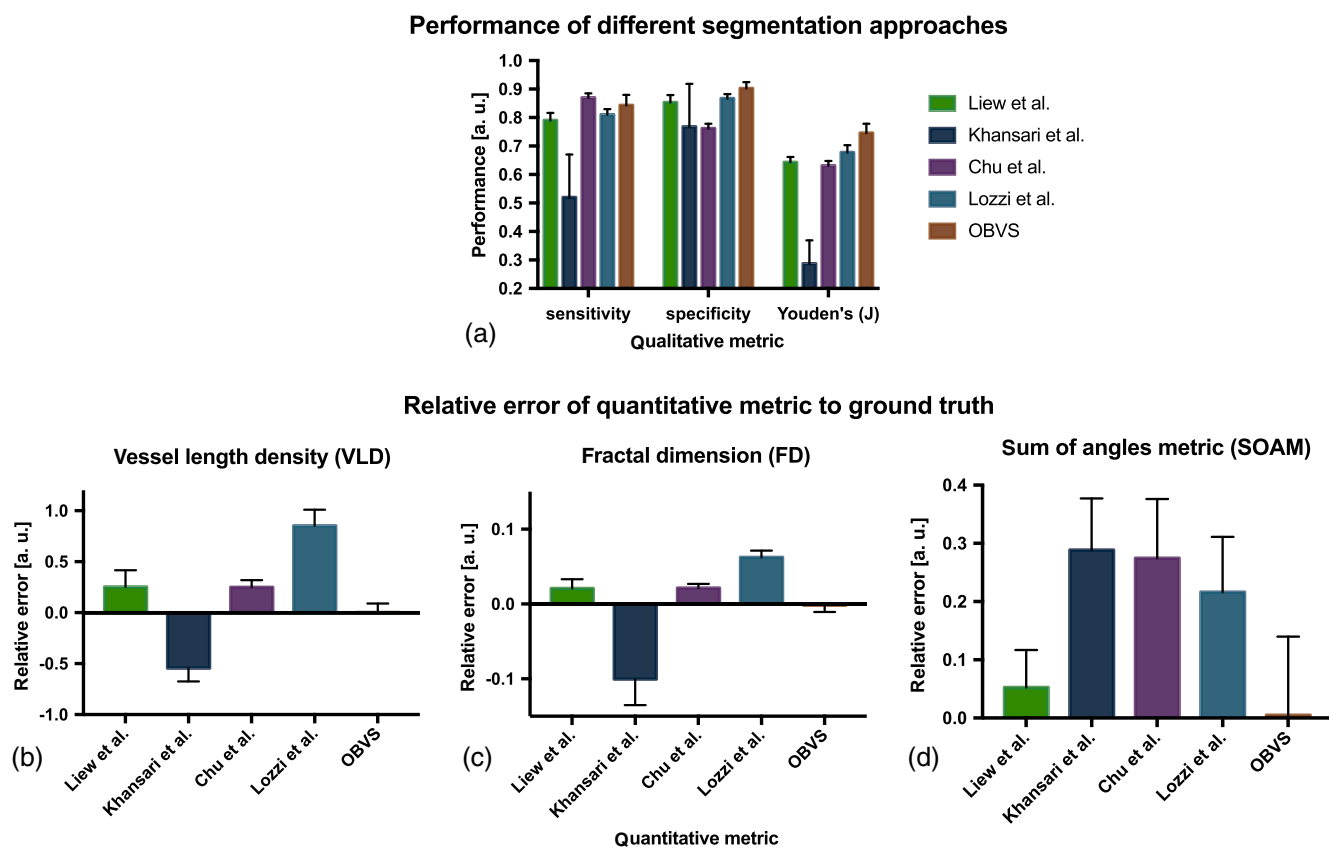
As shown in Figs. 4 and 5, Chu's method[15] had a high sensitivity but poor specificity, generating vessel more in number and caliper than the manual expert could find.

Furthermore, for each method individually, the error of quantitative metrics derived from the vascular network was assessed and given averaged over the five test images in Figs. 6(b)–6(d).

Although this analysis indicates the integrity of each segmentation method, the desired method would produce results with the least difference to the ground truth. As shown in Figs. 6(b) and 6(c), results of OBVS agreed almost completely in terms of VLD, FD, and SOAM with those derived from the ground truth, with 0.1%, 0.2%, and 0.5% errors, respectively.

4 Discussion and Conclusion

The comparison of OBVS with four published algorithms for vessel segmentation in OCTA images demonstrated improved quality and reliability, which was obtained using the pipeline proposed in this work. Although the composition of this pipeline

Performance of different segmentation approaches

(a)

Relative error of quantitative metric to ground truth

(b) (c) (d)

Fig. 6 Performance of the different approaches to OCTA vessel segmentation averaged over the five test images. (a) Congruence with the ground truth (see Sec. 2.4). (b)–(d) Relative difference of quantitative metrics of the vessel architecture to the ground truth (see Sec. 2.5).

was specifically optimized to cope with imaging artifacts, also for uncorrupted images, it showed results superior to the four cited methods. The validity of quantitative metrics assessed for vascular networks was also considerably improved.

Training and testing of OBVS are limited to OCTA images of mouse skin, although we believe that the composition of methods in our OBVS pipeline is applicable for OCTA images of other tissue types such as human skin. The parameters of the methods might need to be adapted.

While creating our pipeline, we found the denoising step to have the most critical impact on the segmentation quality. Despite the fact that the Hessian-based vesselness filter was specifically designed to enhance vascular structures in angiograms,[31] in our investigations it performed with surprisingly poor specificity and thus scored a poor Youden's index J. It adopted imaging artifacts creating phantom vessels in noisy regions (similar to the results of the method of Chu et al. shown in Fig. 4 and upper tile of Fig. 5) and produced overly dilated vessel. The BM3D filter, which had not been applied to OCTA images in prior work, showed superior performance for both sensitivity and specificity. It compensated the noise well, while producing sharp and smooth vessel boundaries and enhancing the contrast between vessel and background.

The procedure to find the optimal segmentation pipeline was inspired by the machine learning algorithms. In contrast to deep-learning-based segmentation, here, we applied not a variety generic filters but the ideal combination of established methods with optimal parameters. This tailoring of the pipeline enabled us to train with a set of training data that was much smaller than

those for deep-learning approaches, while containing also images corrupted by artifacts. For example, Prenta et al.[47] used 80 normal OCTA images for a deep-learning-based segmentation and scored an accuracy of 77% to 83%, whereas OBVS scored here with an accuracy of $87.6\% \pm 0.6\%$ for corrupted images. Note, in comparison to deep-learning, the optimization to obtain OBVS was done subsequently and not globally and OBVS was not evaluated on the data of Prenta et al.

Our optimization was based on a metric for segmentation quality and hence strongly depended on the integrity of the ground-truth segmentation. Using the STAPLE method, we were able to combine the votes of four manual segmentations and find a consensus. Yet, we observed that even the single rater was unable to identify every vessel reliably [indicated as yellow and blue regions in Fig. 1(b)]. This might be due to the rater guessing in areas of artifacts and low contrast of the sample OCTA images. Note, manual labeling was performed on the original images prior to any enhancement. Some commercial OCTA devices might only provide images that already underwent image processing and show modified vessel networks. Nonetheless, with $J = 0.75 \pm 0.01$ [Fig. 6(a)] the proposed automatic method approximatively reached the average Youden's index J of the manual segmentations of the rater, whose four segmentations compared to the STAPLE result scored in average with $J = 0.77 \pm 0.01$.

This sample dataset consisted of images with various types of artifacts that might not occur in different imaging situations or different anatomical sites. For example, air bubbles were absent

when imaging the retina in an intact eye.[21] The challenge of segmenting the vessels in these corrupted images might have led to J scores that were lower than those commonly demonstrated in angiographic imaging modalities other than OCTA in both manual and automatic segmentations. For example in fundus photography the work of Zhao et al. scored with $J = 0.68 - 0.76$[17] and Fraz et al. reviewed methods with scores of $J = 0.63 - 0.80$[48] [we converted these J values from given sensitivity and specificity using Eq. (5) in Sec. 2.4]. Those analyses of fundus photography were benchmarked on open-source image databases, such as STARE and DRIVE, with high-quality manual segmentations obtained by many experts on fundus photography images with minimal artifacts. As such a database is not yet available for OCTA images, especially of skin, our scores of qualitative metrics are not directly comparable to work in other fields.

Yet the visual impression and the analysis of the quantitative metrics show the advantages of our method by providing results closer to the ground truth than other published methods, even when optimized in the same fashion. Moreover, this study demonstrates how much the quantitative vascular network characteristics that are derived from an OCTA image vary when different methods for vessel segmentation are applied to the same images [Figs. 6(b)–6(d)]. Consequently, in any work on automatic analysis of OCTA images, the error of the utilized segmentation method should be assessed by comparison to ground truth data and presented to the reader.

As the development and evaluation of image-processing methods are dependent on the quality of the ground truth, open-source databases have been created in many fields of medical imaging. Unfortunately, in OCTA, such a resource is not available, even though a centralized, open-source database would enable to benchmark angiographic algorithms on raw data (see Sec. 2.2) and image processing and quantification on angiographies. This could lead to a common standard methodology to obtain quantitative metrics of OCTA images. Furthermore, open-access OCTA databases could eventually enable the comparison of medical findings among different publications, and researchers would not need to redo redundant parts of studies but could derive cross-wise conclusions from their and other's findings. However, the image acquisition and processing among commercial OCTA devices have not been standardized yet, which leads to images of heterogeneous quality.[49] Hence, the assembly of such a database at this point would be impracticable due to the variability and amount of commercial devices and developments. Furthermore, its applicability to every single device would be limited.

Disclosures

Acknowledgments

References

1. A. J. Pappano and W. G. Wier, *Cardiovascular Physiology*, 10th ed., Elsevier, Philadelphia, Pennsylvania (2013).
2. P. Martin, "Wound healing—aiming for perfect skin regeneration," *Science* **276**(5309), 75–81 (1997).
3. J. K. Barton and S. Stromski, "Flow measurement without phase information in optical coherence tomography images," *Opt. Express* **13**(14), 5234–5239 (2005).
4. A. Mariampillai et al., "Optimized speckle variance OCT imaging of microvasculature," *Opt. Lett.* **35**(8), 1257–1259 (2010).
5. R. Reif et al., "Quantifying optical microangiography images obtained from a spectral domain optical coherence tomography system," *Int. J. Biomed. Imaging* **2012**, 1–11 (2012).
6. A. Zhang et al., "Methods and algorithms for optical coherence tomography-based angiography: a review and comparison," *J. Biomed. Opt.* **20**(10), 100901 (2015).
7. J. Enfield, E. Jonathan, and M. Leahy, "In vivo imaging of the microcirculation of the volar forearm using correlation mapping optical coherence tomography (CMOCT)," *Biomed. Opt. Express* **2**(5), 1184–1193 (2011).
8. A. S. Nam, I. Chico-Calero, and B. J. Vakoc, "Complex differential variance algorithm for optical coherence tomography angiography," *Biomed. Opt. Express* **5**(11), 3822–3832 (2014).
9. L. Guo et al., "Improved motion contrast and processing efficiency in OCT angiography using complex-correlation algorithm," *J. Opt.* **18**(2), 025301 (2016).
10. J. P. Campbell et al., "Detailed vascular anatomy of the human retina by projection-resolved optical coherence tomography angiography," *Sci. Rep.* **7**, 42201 (2017).
11. M. M. Khansari et al., "Method for quantitative assessment of retinal vessel tortuosity in optical coherence tomography angiography applied to sickle cell retinopathy," *Biomed. Opt. Express* **8**(8), 3796–3806 (2017).
12. J. Staal et al., "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging* **23**(4), 501–509 (2004).
13. S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation," *IEEE Trans. Med. Imaging* **23**(7), 903–921 (2004).
14. K. BahadarKhan, A. A. Khaliq, and M. Shahid, "A morphological hessian based approach for retinal blood vessels segmentation and denoising using region based OTSU thresholding," *PLoS One* **11**(9), e0162581 (2016).
15. Z. Chu et al., "Quantitative assessment of the retinal microvasculature using optical coherence tomography angiography," *J. Biomed. Opt.* **21**(6), 066008 (2016).
16. K. D. Toennies, *Guide to Image Analysis—Methods and Algorithms*, 1st ed., Springer, London (2017).
17. Y. Zhao et al., "Retinal vessel segmentation: an efficient graph cut approach with retinex and local phase," *PLoS One* **10**(4), e0127486 (2015).
18. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Addison-Wesley Longman Publishing Co., 1st ed., Boston, Massachusetts (1992).
19. M. Ulrich et al., "Dynamic optical coherence tomography in dermatology," *Dermatology* **232**(3), 298–311 (2016).
20. A. J. Deegan et al., "Optical coherence tomography angiography of normal skin and inflammatory dermatologic conditions," *Lasers Surg. Med.* **50**(3), 183–193 (2018).
21. R. F. Spaide, J. G. Fujimoto, and N. K. Waheed, "Image artifacts in optical coherence angiography," *Retina* **35**(11), 2163–2180 (2015).
22. Y. M. Liew et al., "In vivo assessment of human burn scars through automated quantification of vascularity using optical coherence tomography," *J. Biomed. Opt.* **18**(6), 061213 (2013).
23. H. H. Carter et al., "Optical coherence tomography in the assessment of acute changes in cutaneous vascular diameter induced by heat stress," *J. Appl. Physiol.* **121**(4), 965–972 (2016).
24. D. Manstein et al., "Fractional photothermolysis: a new concept for cutaneous remodeling using microscopic patterns of thermal injury," *Lasers Surg. Med.* **34**(5), 426–438 (2004).
25. A. Lozzi et al., "Image quality metrics for optical coherence angiography," *Biomed. Opt. Express* **6**(7), 2435–2447 (2015).
26. P. A. Yushkevich et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage* **31**(3), 1116–1128 (2006).
27. W. J. Youden, "Index for rating diagnostic tests," *Cancer* **3**(1), 32–35 (1950).
28. R. Lopes and N. Betrouni, "Fractal and multifractal analysis: a review," *Med. Image Anal.* **13**(4), 634–649 (2009).

29. E. Bullitt et al., "Measuring tortuosity of the intracerebral vasculature from MRA images," *IEEE Trans. Med. Imaging* **22**(9), 1163–1171 (2003).
30. R. Ogniewicz and M. Ilg, "Voronoi skeletons: theory and applications," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, pp. 63–69 (1992).
31. A. F. Frangi et al., "Multiscale vessel enhancement filtering," *Lect. Notes Comput. Sci.* **1496**, 130–137 (1998).
32. W. Niblack, *An Introduction to Digital Image Processing*, Prentice-Hall, Englewood Cliffs, New Jersey (1986).
33. L. Liu et al., "Automated choroidal neovascularization detection algorithm for optical coherence tomography angiography," *Biomed. Opt. Express* **6**(9), 3564–3576 (2015).
34. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–313 (1965).
35. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vision*, Bombay, India, Vol. 6, pp. 839–846 (1998).
36. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990).
37. A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit.*, Vol. 2, pp. 60–65 (2005).
38. K. Dabov et al., "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
39. J. Frankle and J. McCann, "Method and apparatus for lightness imaging," U. S. Patent No. 4, 384, 336 (1983).
40. G. Lathen, J. Jonasson, and M. Borga, "Blood vessel segmentation using multi-scale quadrature filtering," *Pattern Recognit. Lett.* **31**(8), 762–767 (2010).
41. J. S. Weszka, "A survey of threshold selection techniques," *Comput. Graphics Image Process.* **7**(2), 259–265 (1978).
42. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
43. T. W. Ridler and S. Clavard, "Picture thresholding using an iterative selection method," *IEEE Trans. Syst. Man Cybern.* **8**(8), 630–632 (1978).
44. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001).
45. Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(9), 1124–1137 (2004).
46. R. Strzodka and A. Telea, "Generalized distance transforms and skeletons in graphics hardware," in *Proc. EG/EEE TCVG Symp. Visualization (VisSym)*, pp. 221–230 (2004).
47. P. Prenta et al., "Segmentation of the foveal microvasculature using deep learning networks deep learning networks," *J. Biomed. Opt.* **21**(7), 075008 (2016).
48. M. M. Fraz et al., "Blood vessel segmentation methodologies in retinal images—a survey," *Comput. Meth. Programs Biomed.* **108**(1), 407–433 (2012).
49. M. R. Munk et al., "Oct-angiography: a qualitative and quantitative comparison of 4 OCT-A devices," *PLoS One* **12**(5), e0177059 (2017).

Biographies of the authors are not available.