ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Original Article

Deep learning differentiates between healthy and diabetic mouse ears from optical coherence tomography angiography images

Martin Pfister,[1,2,3] (ID) Hannes Stegmann,[1,2] Kornelia Schützenberger,[1,2] Bhavapriya Jasmin Schäfer,[1,2] Christine Hohenadl,[2,4] Leopold Schmetterer,[1,2,5,6,7,8,9] Martin Gröschl,[3] and René M. Werkmeister[1,2] (ID)

[1]Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria. [2]Christian Doppler Laboratory for Ocular and Dermal Effects of Thiomers, Medical University of Vienna, Vienna, Austria. [3]Institute of Applied Physics, Vienna University of Technology, Vienna, Austria. [4]Croma Pharma GmbH, Leobendorf, Austria. [5]Department of Clinical Pharmacology, Medical University of Vienna, Vienna, Austria. [6]Singapore Eye Research Institute, Singapore. [7]School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore. [8]Ophthalmology and Visual Sciences Academic Clinical Program, Duke-NUS Medical School, Singapore. [9]Institute of Molecular and Clinical Ophthalmology, Basel, Switzerland

Address for correspondence: René M. Werkmeister, Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Währinger Gürtel 18–20/4L, 1090 Vienna, Austria. rene.werkmeister@meduniwien.ac.at

We trained a deep learning algorithm to use skin optical coherence tomography (OCT) angiograms to differentiate between healthy and type 2 diabetic mice. OCT angiograms were acquired with a custom-built OCT system based on an akinetic swept laser at 1322 nm with a lateral resolution of ∼13 μm and using split-spectrum amplitude decorrelation. Our data set consisted of 24 stitched angiograms of the full ear, with a size of approximately 8.2 × 8.2 mm, evenly distributed between healthy and diabetic mice. The deep learning classification algorithm uses the ResNet v2 convolutional neural network architecture and was trained on small patches extracted from the full ear angiograms. For individual patches, we obtained a cross-validated accuracy of 0.925 and an area under the receiver operating characteristic curve (ROC AUC) of 0.974. Averaging over multiple patches extracted from each ear resulted in the correct classification of all 24 ears.

Keywords: angiographic imaging; diabetes; machine learning; optical coherence tomography

Introduction

Diabetes mellitus is a chronic metabolic disorder characterized by elevated levels of blood glucose over a prolonged period. It is linked to macro- and microangiopathies that cause several vascular diseases, such as diabetic retinopathy (DR), diabetic nephropathy, diabetic neuropathy, and arteriosclerosis. More than 425 million people worldwide are currently affected by diabetes, with its prevalence continuing to rise.[1] Diabetes may affect many parts of the body, but changes to the vasculature are best observed in the retina, where they can be examined using noninvasive optical methods, such as fundus photography or optical coherence tomography (OCT).

OCT is a noninvasive imaging modality that can capture three-dimensional images of tissue samples with axial and lateral resolutions of a few micrometers.[2–4] Analogous to ultrasound, OCT measures the magnitude and echo time delay of back-reflected or back-scattered light to produce structural images of the sample. Today, OCT is widely used in ophthalmology, but it also has promising applications in other medical fields, such

as dermatology, where it can provide cross-sectional images revealing skin morphology down to a depth of a few millimeters. A number of functional extensions to OCT provide access to secondary information in addition to the morphological imaging offered by standard OCT. In particular, optical coherence tomography angiography (OCTA) enables visualization of the vascular network of the probed tissue without the need for a contrast agent.[5]

The advent of deep learning, spurred by the recent availability of large data sets and sufficient computing power, has triggered tremendous advances in computer vision.[6,7] Naturally, this new technique has been enthusiastically adopted in the field of medical imaging, with hundreds of papers published already.[8–10] Retinal fundus photographs have been extensively used to train machine learning algorithms to diagnose diabetes-associated conditions with excellent results; for example, Gulshan *et al.*[11] demonstrated a deep learning algorithm for detecting DR in retinal fundus photographs that had an area under the receiver operating characteristic curve (ROC AUC) of 0.991 and 0.990 for two different validation data sets. Visualization of other pathologies, such as diabetic macular edema (DME), is improved in retinal OCT images as compared with fundus photographs, because OCT also provides additional depth information. For example, Kermany *et al.*[12] trained a deep learning system on OCT images to screen patients for retinal diseases and obtained an ROC AUC of 0.9987 for DME versus normal. Both fundus photography and OCT imaging present features of the retinal vasculature—the first in an *en face* manner, the latter as cross-sectional data. Since OCT angiography has already been shown to be superior to fundus photography for detection of vascular abnormalities in DR when assessed manually,[13,14] it should also be well suited for the detection of diabetic vascular changes by machine learning. In a very recent paper, Sandhu *et al.*[15] used clinical biomarkers and quantitative parameters derived from OCT and OCTA images to train a random forest classifier to diagnose nonproliferative DR with an ROC AUC of 0.981 for DR versus no DR. If limited to only the OCTA-derived parameters, they were still able to obtain an ROC AUC of 0.937. Additionally, much research has been performed to improve the performance of machine learning algorithms on public retinal fundus photograph databases. For example, Wang *et al.*[16] aimed

to mimic the zoom-in process of a clinician to examine retinal images in their Zoom-in-Net; DME-Net by He *et al.*[17] added a segmentation step before classifying for DME; and CANet by Li *et al.*[18] used disease-specific attention modules and exploited the correlation between DR and DME to improve performance.

In the present study, we employed OCT for noninvasive angiography imaging in an *in vivo* model in the pinna of healthy and type 2 diabetic mice. The primary aim was to investigate whether a deep learning algorithm can differentiate between the vasculatures of nondiabetic and diabetic animals on the basis of OCT angiograms. In an effort to evaluate differences in the images that might lead to the recognition by the deep learning algorithm, we performed further analyses, including comparison of the OCTA parameters vessel density, length, diameter, and number of bifurcations as determined from the angiographic images.

Materials and methods

Animals
Six female healthy (C57BL/6J) and six female type 2 diabetic mice (BKS.Cg-Dock7m+/+ Leprdb/J) were included in this study. The animals were purchased from the Charles River Laboratory and kept under controlled, standardized conditions (artificial light/dark cycle of 12:12, room temperature at 22 ± 2 °C, humidity at 45 ± 10%) at the Center for Biomedical Research of the Medical University of Vienna. The animals had *ad libitum* access to water and complete feed for mice (sniff R/M-H, sniff Spezialdiäten GmbH, Soest, Germany). All diabetic mice had a blood sugar level greater than 300 mg/dL at the time of the study. OCT angiography images were acquired from the mice at the age of 10 weeks while they were under anesthesia from breathing 2% isoflurane. OCTA images were recorded at the pinna, where, owing to its flat shape, most vessels are located close to the skin surface and thus within the accessible penetration depth and optimal focal range of the OCT system. Mouse ears were depilated before the measurements in order to avoid shadowing by body hair. The study was approved by the local Animal Welfare Committee of the Medical University of Vienna and the Austrian Federal Ministry of Education, Science, and Research and was fully compliant with Austrian legislation.

Table 1. Cross-validation results per patch and per image

	Accuracy per patch	ROC per patch	Accuracy per image	ROC per image
Total	0.925	0.974	1.000	1.000
Fold 1	0.907	0.983	1.000	1.000
Fold 2	0.913	0.984	1.000	1.000
Fold 3	0.980	0.999	1.000	1.000
Fold 4	0.908	0.972	1.000	1.000
Fold 5	0.982	0.998	1.000	1.000
Fold 6	0.865	0.899	1.000	1.000

NOTE: Each fold consists of many small, overlapping patches extracted from two nondiabetic and two diabetic whole-ear OCTA images. Results are given both for individual patches and whole images (i.e., using the averages over all patches of each image).

Image acquisition and data set

OCT images were acquired using a custom-built OCT prototype. The system is based on an akinetic swept laser with a flat spectrum centered at 1322 nm and a bandwidth of 62 nanometers. A detailed description of this system has been published previously.[19] For the current angiography measurements, a scan lens with a focal length of 18 mm (Thorlabs LSM02, Thorlabs GmbH, Bergkirchen, Germany), providing a lateral resolution of ∼13 μm with a depth of focus of ∼200 μm, was used. Image acquisition was performed over a square skin patch of 4.86 × 4.86 millimeters. Angiography images were generated using split-spectrum amplitude decorrelation[20] from four B-scans recorded at each position. After applying the thresholded intensity data as a mask to the angiography data, a 3D Gaussian filter with a small sigma was used to reduce speckle noise. *En face* images were generated by applying maximum projection in axial direction and had a size of 768 × 768 pixels. Four images were taken from slightly overlapping regions of the mouse ear and stitched together, resulting in large angiograms with a size of approximately 8.2 × 8.2 mm, covering nearly the entire pinna. The final data set consisted of a total of 24 full-ear angiograms evenly distributed between healthy and diabetic mice.

Deep learning algorithm

A convolutional neural network (CNN) was trained to distinguish OCTA images of healthy and diabetic mice. The network uses a 50-layer ResNet v2[21] with a two-way fully connected layer as the final layer. The network was initialized with weights pretrained on ImageNet and all layers were fully retrained. Training was performed for 30,000 iterations using cross-entropy loss and stochastic gradient descend with a learning rate of 0.0001, Nesterov momentum of 0.9, and a batch size of 128. To limit overfitting, L_2 regularization was applied. Instead of using the full OCTA images as input, patches of 330 × 330 pixels (2.09 × 2.09 mm) were extracted and resized to the network input size of 224 × 224 pixels. This helped to generate a sufficiently large training set from the available images. Only patches that did not extend beyond the edge of the pinna were used for training. To introduce further variability into the training set, extensive image augmentation was applied, including random contrast and brightness adjustments, mirroring, rotation, and random elastic distortion using a 5 × 5 displacement grid.[22] We employed six-fold cross-validation, with each of the six folds consisting of the angiograms of the ears of one healthy and one diabetic mouse. To measure the performance of the deep learning algorithm, we determined the accuracy and the ROC AUC both for individual patches and after averaging over each mouse ear. The ROC curve plots the true positive rate against the false positive rate for various thresholds. The area under this curve (ROC AUC) is a performance measure of a binary classifier, with an ROC AUC value of 1.0 corresponding to perfect results and 0.5 indicating equal to random performance of the classifier. The CNN was implemented with the help of the Keras library[23] using the TensorFlow backend.[24] Training of one network took approximately 3.3 h using two Nvidia GeForce GTX 1080 Ti GPUs.

Saliency maps

While deep learning algorithms are frequently able to produce astonishingly accurate predictions, their inner workings often remain opaque. One common method of CNN analysis is attention mapping. Of the different saliency mapping techniques that have been discussed in the literature, the most straightforward method is to calculate the derivative (i.e., gradient) of the prediction $S(x)$ with respect to the input image x.[25–27] The gradient $\partial S(x)/\partial x$ quantifies how much a change in a pixel of the input image would change the prediction. While gradient saliency maps have been proven to be rather robust,[28] they are very noisy. SmoothGrad[29] seeks to alleviate this by averaging over multiple saliency maps of noisy copies of the input image. Here, we use a slightly different technique: as our CNN input is extracted from larger angiograms, we can slightly shift the input region and thus average over multiple saliency maps for the same pixel positions in the larger angiography image. This produces large-scale saliency maps for areas greater than the CNN input size, with considerably reduced noise.

Vascular network analysis

In order to investigate the differences in the vasculature of healthy and diabetic mice as revealed by OCTA, four parameters of the vessel networks were determined:[30] First, the vessel density, that is, the percentage of pixels in the image that show vessels, was calculated. This was done by first applying a Sato filter of curvilinear structures[31] to the two-dimensional OCTA images, followed by contrast limited adaptive histogram equalization.[32] Thresholding using Otsu's method[33] produces a binary image, of which the percentage of foreground pixels is the vessel density. Next, the vessel length normalized to a fixed area of 1 mm^2 and calculated from the skeletonized[34] binary image, and the mean vessel diameter, which is the ratio between the vessel density and the normalized vessel length, was determined. These parameters were calculated in an effort to investigate if a possible difference in vessel density may be caused by either more or thicker vessels. Finally, the fourth parameter is the number of bifurcations normalized to vessel length. The four parameters were calculated using the scikit-image[35] and skan[36] libraries.

Independent two-sample t-tests were performed to compare the vascular parameters of the dia-
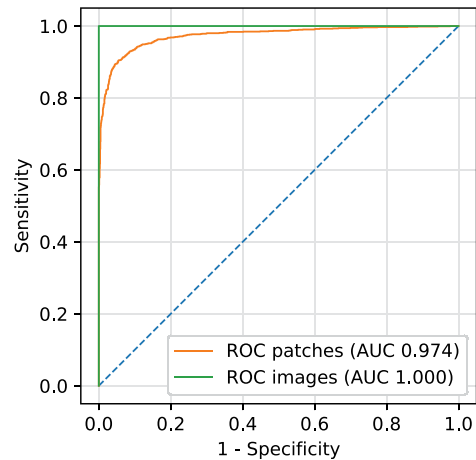
Figure 1. Receiver operating characteristic for the detection of diabetes. ROC patches give the ROC curve for individual overlapping patches extracted from the acquired OCT images. ROC images give the ROC curve after averaging over all usable patches of each image.

betic and healthy mice. Statistical significance was accepted at $P < 0.05$.

Results

Deep learning results

Since the input size for the neural network is smaller than the size of the acquired OCTA images, we extracted 196 overlapping patches from each image using an equally spaced 14 × 14 grid for evaluation of the network performance. On average, adjacent patches overlapped by ∼77%. Only patches that did not extend beyond the edge of the pinna were included in the evaluation data set. In total, we evaluated the six networks on 3026 patches and achieved an accuracy of 0.925 and an ROC AUC of 0.974. When averaging the results of all usable patches per ear (i.e., those not extending beyond the edge of each ear), prediction was correct for all 24 ears (accuracy of 1.0 and ROC AUC of 1.0). Table 1 gives the accuracy and ROC AUC for each of the six networks trained for cross-validation. Figure 1 shows the ROC curves for both kinds of evaluations. Exemplary predictions of the CNN are given in Figures 2 and 3. Figure 4 shows the angiograms and CNN predictions of the healthy and diabetic mouse where the performance of the neural network was worst, both of which still produced a correct prediction after averaging over all usable patches. A few angiograms revealed slight movement artifacts
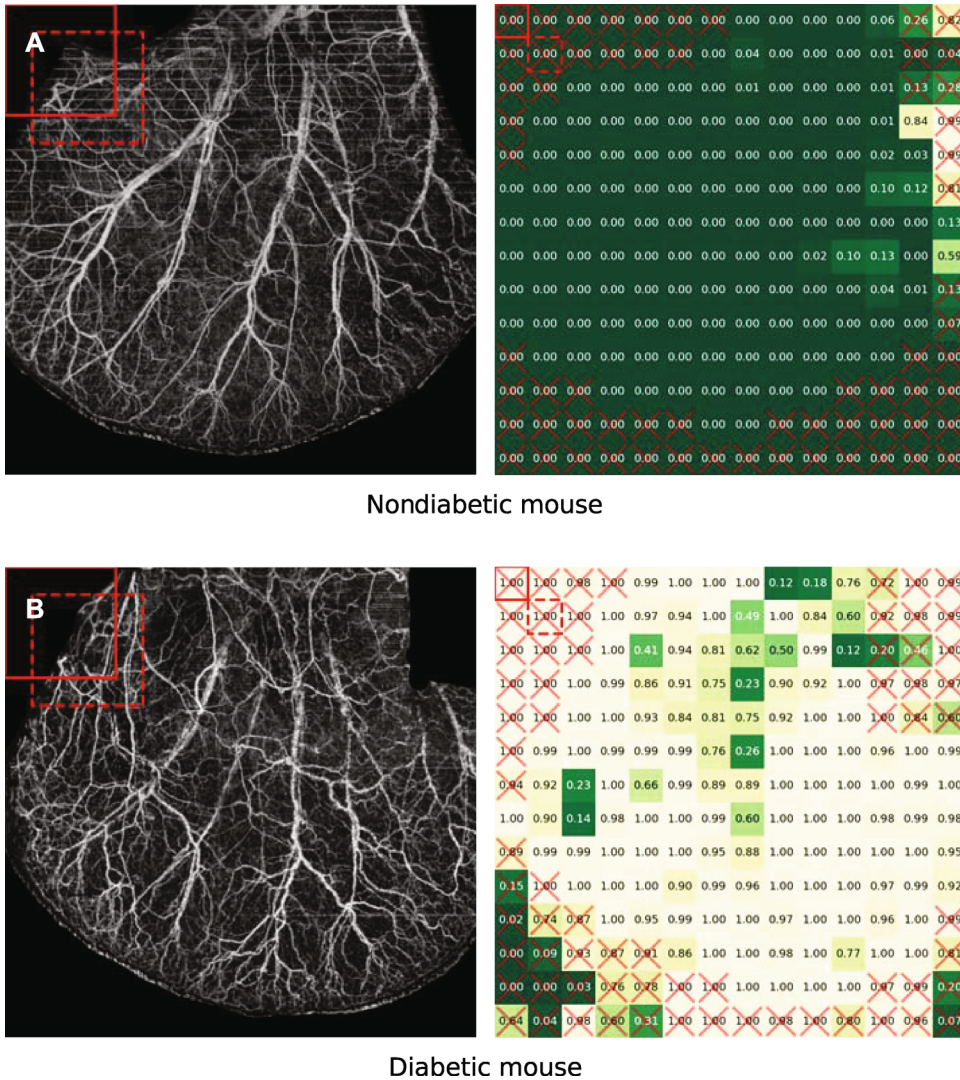
Nondiabetic mouse

Diabetic mouse
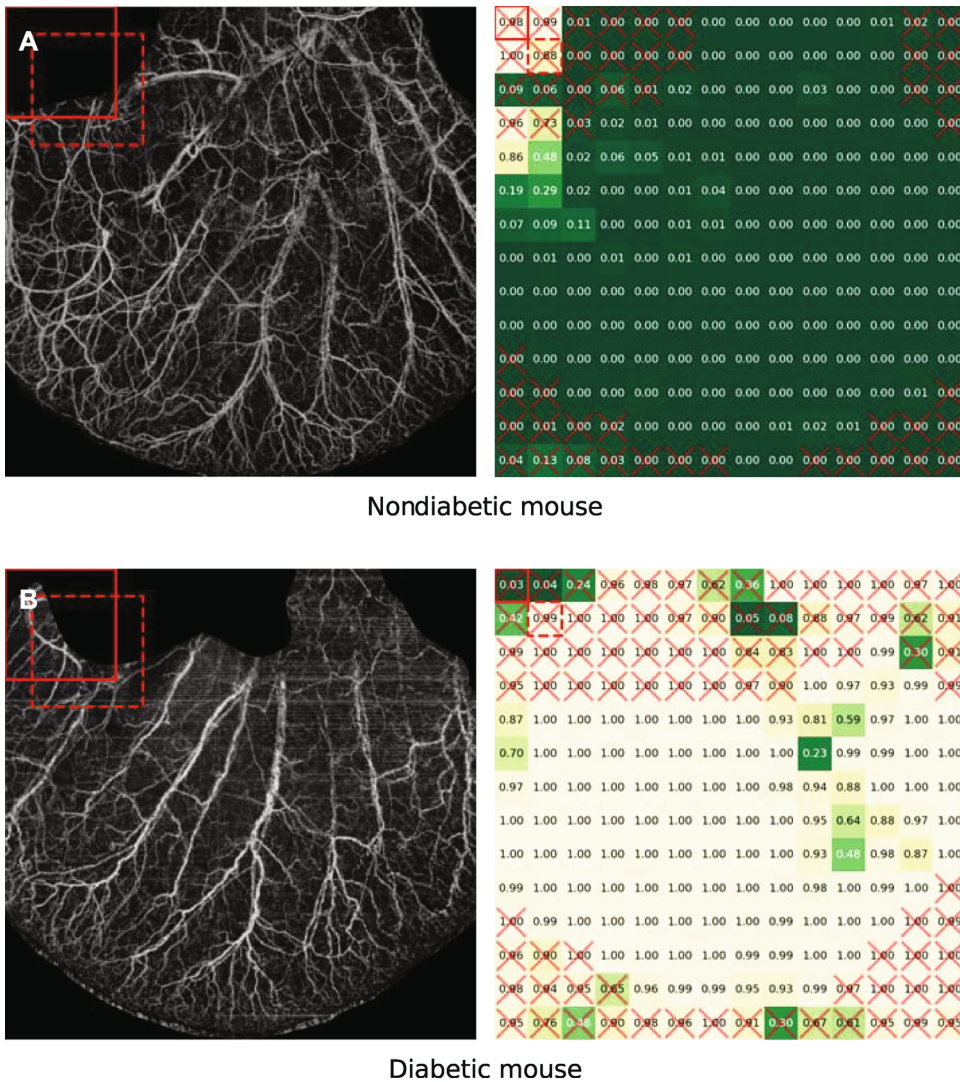
Figure 2. Example of OCTA images with a size of approximately 8.2 × 8.2 mm (left) with corresponding predictions of the CNN (right). As the network expects an input size of 2.09 × 2.09 mm (indicated by the red squares), inference was performed 196 times per image on overlapping patches of this size. Patches extending beyond the useful angiography image area were excluded from the average (indicated by red crosses). Numbers close to 1.0 represent CNN prediction of diabetic vascular changes. (A) Nondiabetic mouse. Average prediction 0.02; median prediction 0.00. (B) Diabetic mouse. Average prediction 0.97; median prediction 1.00.

owing to breathing of the mice (e.g., Figs. 2B and 3A).

Analysis of the deep learning algorithm
Gradient saliency maps for the central regions of a healthy and diabetic mouse ear are given in Figure 5. The gradient is small at the location of larger vessels and large in those areas of the angiogram where no distinct vessels are visually apparent.

To obtain a better understanding on the importance of high-resolution spatial features, we also trained and evaluated CNNs on blurred copies of the angiograms. The results of this analysis are shown in Figure 6. As expected, larger Gauss kernels have a higher impact on performance, but even a Gaussian filter with only $\sigma = 2$ already affects performance noticably, with the cross-validated ROC AUC for individual image patches dropping from 0.974 to 0.933. Still, for a Gauss kernel

Nondiabetic mouse

Diabetic mouse

Figure 3. Example of OCTA images with a size of approximately 8.2 × 8.2 mm (left) with corresponding predictions of the CNN (right). As the network expects an input size of 2.09 × 2.09 mm (indicated by the red squares), inference was performed 196 times per image on overlapping patches of this size. Patches extending beyond the useful angiography image area were excluded from the average (indicated by red crosses). Numbers close to 1.0 represent CNN prediction of diabetic vascular changes. (A) Nondiabetic mouse. Average prediction 0.02; median prediction 0.00. (B) Diabetic mouse. Average prediction 0.90; median prediction 0.99.

with σ = 8 that suppresses all capillaries and leaves only the midsized and larger vessels visible to the observer (Fig. 6E), the network still performs better than random, with an ROC AUC of 0.694.

Finally, we aimed to investigate whether the network also relies on the intensity distribution in the angiograms for the classification. To this end, we performed training and evaluation on permuted copies of the images that only retained the pixel intensity values, but were stripped of any spatial information (see Fig. 6G for an example). If there were no correlation between intensity distribution and diabetes status of the mice, an ROC AUC of 0.5 would be expected, but we obtained an ROC AUC of 0.708, which supports our assumption.

Vascular network analysis
Analysis of the vascular network was performed by calculating four parameters from binarized angiography images: vessel density in percent of pixels,

Nondiabetic mouse

Diabetic mouse
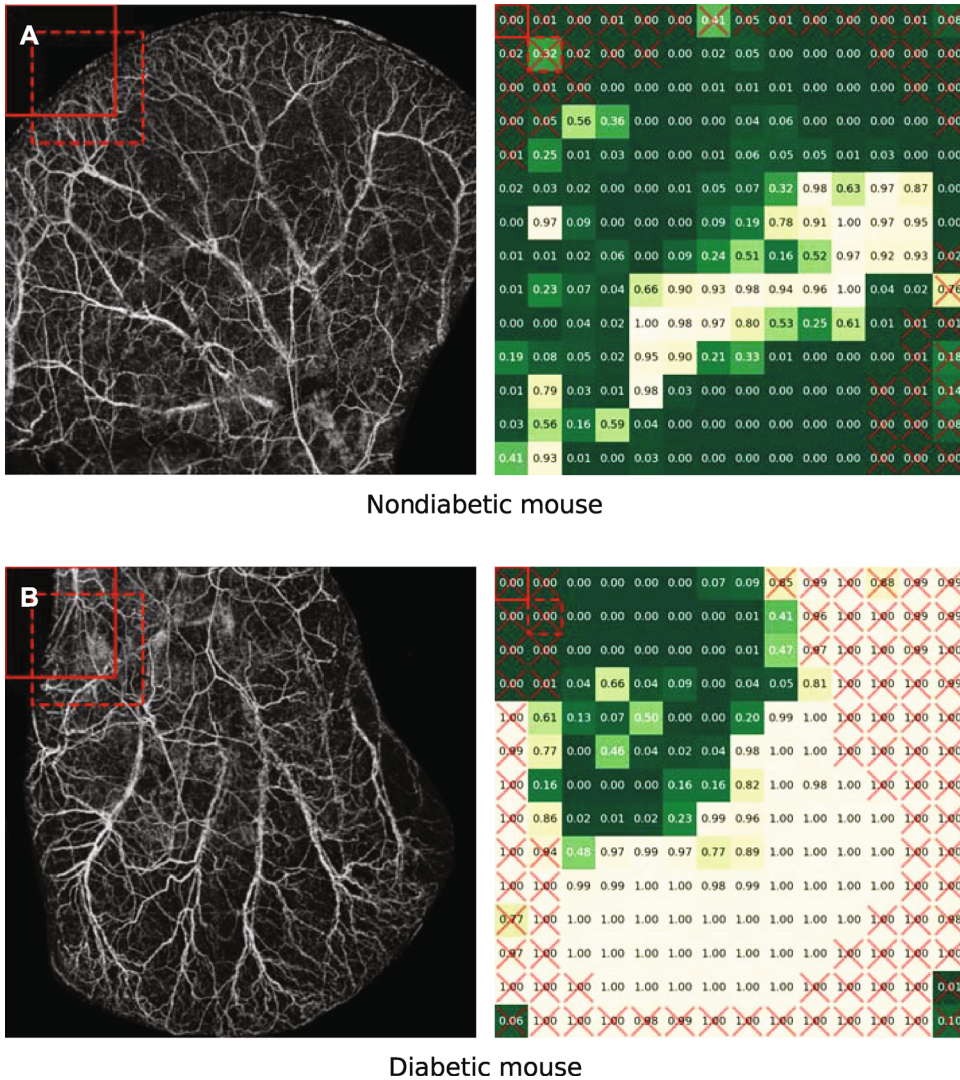
Figure 4. Example of OCTA images with a size of approximately 8.2 × 8.2 mm (left) with corresponding predictions of the CNN (right). As the network expects an input size of 2.09 × 2.09 mm (indicated by the red squares), inference was performed 196 times per image on overlapping patches of this size. Patches extending beyond the useful angiography image area were excluded from the average (indicated by red crosses). Numbers close to 1.0 represent CNN prediction of diabetic vascular changes. The two angiograms shown in this figure are the ones that exhibited the worst performance of the deep learning algorithm. (A) Nondiabetic mouse. Average prediction 0.24; median prediction 0.03. (B) Diabetic mouse. Average prediction 0.58; median prediction 0.88.

normalized vessel length in mm per mm² image area, mean vessel diameter in μm, and number of bifurcations per mm vessel length. Results are given in Table 2 and visualized via box plots given in Figure 7. Both vessel density and normalized vessel length are slightly larger for healthy mice than for diabetic mice, though the difference is not significant ($P = 0.19$ for both). Mean vessel diameter is almost identical ($P = 0.86$) between healthy and diabetic animals. Vasculature in healthy mice exhibits more bifurcations per unit vessel length, but the difference is again not significant ($P = 0.18$).

In total, all four investigated parameters did not show any significant differences that could serve as a distinguishing feature between healthy and diabetic samples.
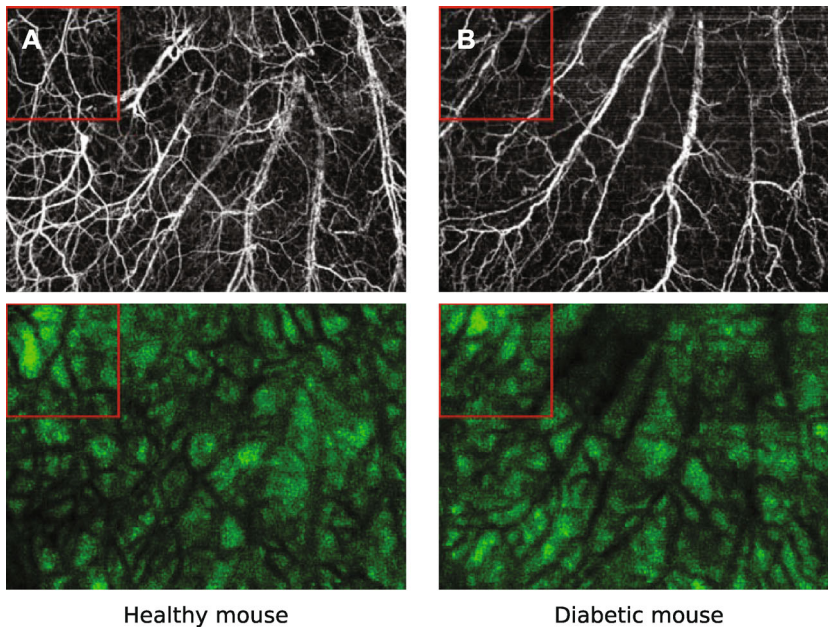
Figure 5. Example of OCTA images (top) and corresponding gradient saliency maps stitched from many overlapping CNN predictions (bottom). Bright green areas represent regions that the network is most influenced by when predicting diabetes-correlated vascular changes. Red squares indicate CNN input size. (A) Healthy mouse. Central region of the ear in Figure 2A. (B) Diabetic mouse. Central region of the ear in Figure 2B.

Discussion

We successfully trained a deep learning algorithm to differentiate between vasculatures of healthy and diabetic mice from OCT angiograms obtained from ears.

As input, small extracted patches rather than the full angiograms were used to train the CNN. While it would be possible to feed a CNN with angiograms of the full ear, using only these small patches has two major advantages: first, deep learning generally needs vast amounts of training images to be effective. By extracting small patches from the full image, hundreds of different training samples can be generated from each mouse. If image augmentation is employed, the number of samples can be even further increased. Second, training a CNN on small images reduces the number of necessary computations per convolution and thus speeds up training. The main disadvantage of this approach is the limited receptive area of the CNN, which probably leads to a higher error rate than if the CNN had access to the full image. We compensated for this by applying the CNN on different patches extracted from each full ear image and averaging the predictions of the CNN. With this approach, we obtained correct classification results for all 24 ears, that is, both a cross-validated accuracy and an ROC AUC of 1.0.

Table 2. Analysis of the vascular network in the ears of healthy (BL6) and diabetic (db/db) mice

	BL6	db/db	*P* value
Vessel density (%)	35.33 (3.02)	33.59 (2.97)	0.19
Normalized vessel length (mm/mm^2)	7.62 (0.64)	7.23 (0.70)	0.19
Mean vessel diameter (μm)	46.39 (1.43)	46.50 (1.59)	0.86
Bifurcations per vessel length (1/mm)	5.02 (0.33)	4.84 (0.26)	0.18

NOTE: Given are the mean values with the standard deviations in parentheses.
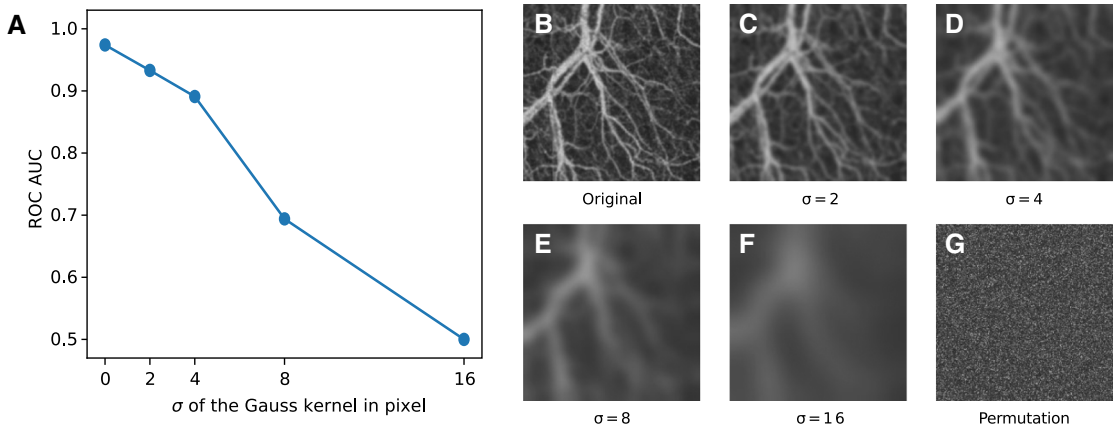
Figure 6. Effects of removing image information on the prediction of diabetes. (A) Impaired prediction due to the removal of high-frequency information with Gaussian blurring. (B) Sample CNN input image. (C–F) Same image with varying degrees of Gaussian blur applied. (G) Same image with randomly shuffled pixel positions. For the permuted images, an ROC AUC of 0.708 was obtained.

In addition to the deep learning analysis, we determined four quantitative vascular parameters in order to investigate if they can explain the nature of the differences between the two groups that the deep learning algorithm was able to exploit. Immunohistochemistry of the skin in diabetic subjects has shown not only inflammation but also an increased blood vessel density, even though skin blood flow was not significantly different compared with non-diabetic subjects.[37] We found, however, that none of the differences in OCTA-based vessel density, mean vessel diameter, normalized vessel length, and number of bifurcations were statistically significant. The deep learning algorithm thus seems to include other image features not reflected in OCTA parameters obtained after postprocessing of the angiograms to achieve its high classification accuracy.

In order to evaluate the operating mode of the CNN, we performed several additional analyses. We calculated gradient saliency maps and found that the maximum attention occurs not at the location of the larger vessels, but rather in the low-contrast areas in between. This suggests that the CNN tends not to look at these large vessels, but rather focuses on the areas that only contain capillaries, and is in agreement with the literature describing diabetes-related damage to small vessels. For example, in the retina, patients with early stages of type 2 diabetes, but without any signs of DR, showed disruptions in the parafoveal capillary network,[38] and patients with DR showed anomalies

in the foveal microvasculature.[39] In the skin, capillary circulation in the toes of patients with diabetes was severely reduced, even though the total skin microcirculation was normal,[40] and patients with diabetic foot ulcer had fewer capillaries in the dermal papillary layer and showed detrimental remodeling.[41]

When training and evaluating CNNs on blurred copies of the angiograms, we found that even a mild blurring already deteriorates the prediction accuracy, but performance remained better than average up to fairly strong blurring with a Gauss kernel of $\sigma = 8$, which effectively eliminates all but the largest vessels. Thus, both high- and low-frequency spatial information seem to play an important role in the prediction of diabetic vascular changes. Finally, we trained and evaluated CNNs on permuted copies of the images and obtained an accuracy that was better than random, indicating that the CNN is able to exploit a difference in the intensity distributions in the angiograms of the two groups. This suggests that some of the differences between the vascular systems of healthy and diabetic mice might be independent of the shape of the vessels, but rather related to the decorrelated pixels and thus to the perfusion of the existing vessels, which again corresponds to the literature cited above.

A few limitations have to be considered when interpreting the results of our study: First, small capillaries have a diameter below the resolution limit of our OCT system. While they might still
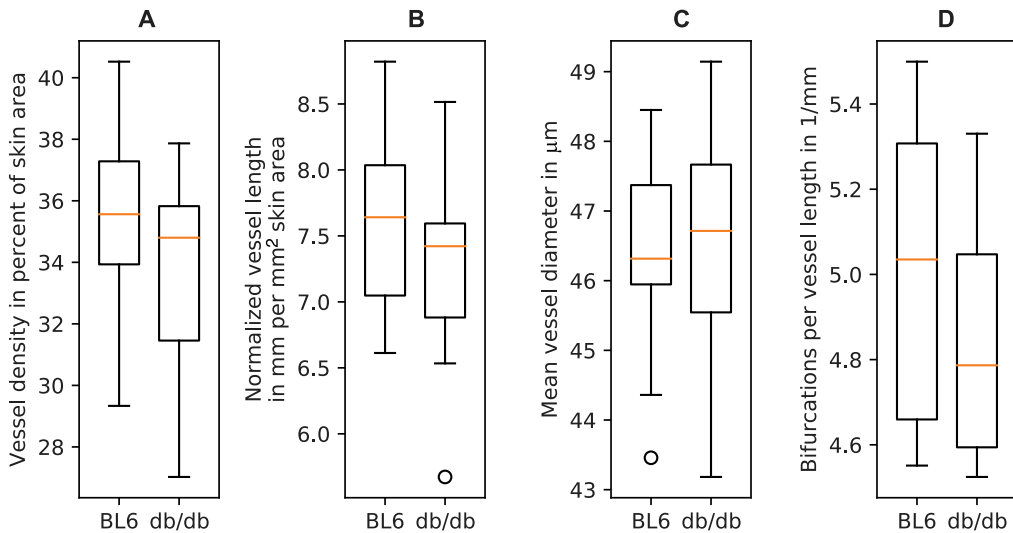
Figure 7. Analysis of the vascular network in the ears of healthy (BL6) and diabetic (db/db) mice. The box plots show (A) vessel density, (B) normalized vessel length, (C) mean vessel diameter, and (D) bifurcations per vessel length.

appear as nonconnected decorrelated regions in the angiograms and thus may be considered by the CNN, they will not be included in the deduced quantitative parameters. Second, we performed the analysis of vascular parameters on the same two-dimensional images that were also used for training the CNN. As these 2D images are maximum intensity projections of the volumetric data set, they do not convey the depth position of the vessels, possibly leading to extra bifurcations that do not exist in the proper three-dimensional vascular system. A three-dimensional analysis of the volumetric OCT angiograms would avoid this, but since the main focus of this study is the implementation of a deep learning algorithm for image classification, we chose to use the same 2D image for both methods in order to be able to compare the results better. Third, our data set is limited in size and contains only 24 full ear angiograms. Still, as our approach fundamentally works with small image patches instead of full ear angiograms, we are confident in the significance of our results for the evaluation of 3026 patches that showed a cross-validated ROC AUC of 0.974. Fourth, as a consequence of the disease, the diabetic mice are obese. Changes in the vasculature observed by the CNN could also be due to the obesity and thus only indirectly related to diabetes. Additional experiments would be needed to study this issue, for example, comparing obese mice suffering from diabetes with mice that develop obe-

sity because of a lack of sensation of satiety. Finally, regarding the applicability in humans, OCT angiography in human skin is possible, but its greater thickness might make the analysis more difficult. Should analysis of diabetes-induced vascular damages from human skin OCT angiograms be found to be impracticable, the human retina may serve as an alternative window into the state of vascular health in humans.

Our study demonstrates that a deep learning algorithm can successfully detect diabetes-induced changes in the vascular system, warranting further research in this area. Here, future studies should compare angiograms of different stages of diabetes. Deep learning analysis of OCT angiography images has profound potential to offer new biomarkers that help with early detection and treatment of the disease, ultimately preserving quality of life for patients affected by diabetes.

Acknowledgments

Author contributions

M.P., C.H., L.S., M.G., and R.M.W. contributed to the conception and design of the work. M.P.,

K.S., and B.J.S. acquired the data. M.P., H.S., and R.M.W. conducted the analyses and interpreted the results. M.P. and R.M.W. wrote the initial draft. All authors participated in revising the manuscript and approved the submitted version. M.P. accepts responsibility for the integrity of the data analyzed.

Competing interests

The authors declare no competing interests.

References

1. International Diabetes Federation. 2017. *IDF Diabetes Atlas*. 8th edn. Brussels: International Diabetes Federation.
2. Huang, D., E.A. Swanson, C.P. Lin, *et al.* 1991. Optical coherence tomography. *Science* **254:** 1178–1181.
3. Swanson, E.A., J.A. Izatt, M.R. Hee, *et al.* 1993. *In vivo* retinal imaging by optical coherence tomography. *Opt. Lett.* **18:** 1864–1866.
4. Fercher, A.F., C.K. Hitzenberger, W. Drexler, *et al.* 1993. *In vivo* optical coherence tomography. *Am. J. Ophthalmol.* **116:** 113–114.
5. Makita, S., Y. Hong, M. Yamanari, *et al.* 2006. Optical coherence angiography. *Opt. Express* **14:** 7821–7840.
6. Krizhevsky, A., I. Sutskever & G.E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. In *Proceedings of Advances in Neural Information Processing Systems*, 1090–1098.
7. LeCun, Y., Y. Bengio & G. Hinton. 2015. Deep learning. *Nature* **521:** 436–444.
8. Litjens, G., T. Kooi, B.E. Bejnordi, *et al.* 2017. A survey on deep learning in medical image analysis. *Med. Image Anal.* **42:** 60–88.
9. Schmidt-Erfurth, U., A. Sadeghipour, B.S. Gerendas, *et al.* 2018. Artificial intelligence in retina. *Prog. Retin. Eye Res.* **67:** 1–29.
10. Soffer, S., A. Ben-Cohen, O. Shimon, *et al.* 2019. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology* **290:** 590–606.
11. Gulshan, V., L. Peng, M. Coram, *et al.* 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316.** 2402.
12. Kermany, D.S., M. Goldbaum, W. Cai, *et al.* 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **172:** 1122–1131.e9.
13. Khadamy, J., K. Aghdam & K. Falavarjani. 2018. An update on optical coherence tomography angiography in diabetic retinopathy. *J. Ophthalmic Vis. Res.* **13:** 487–497.
14. Schaal, K.B., M.R. Munk, I. Wyssmueller, *et al.* 2019. Vascular abnormalities in diabetic retinopathy assessed with swept-source optical coherence tomography angiography widefield imaging. *Retina* **39:** 79–87.
15. Sandhu, H.S., M. Elmogy, A. Taher Sharafeldeen, *et al.* 2020. Automated diagnosis of diabetic retinopathy using clinical biomarkers, optical coherence tomography, and optical coherence tomography angiography. *Am. J. Ophthalmol.* **216:** 201–206.

16. Wang, Z., Y. Yin, J. Shi, *et al.* 2017. Zoom-in-Net: deep mining lesions for diabetic retinopathy detection. In *Medical Image Computing and Computer Assisted Intervention - MICCAI 2017.* M. Descoteaux, L. Maier-Hein, A. Franz *et al.*, Eds.: 267–275. Cham: Springer International Publishing.
17. He, X., Y. Zhou, B. Wang, *et al.* 2019. DME-Net: diabetic macular edema grading by auxiliary task learning. In *Medical Image Computing and Computer Assisted Intervention MICCAI 2019.* D. Shen, T. Liu, T.M. Peters *et al.*, Eds.: 788–796. Cham: Springer International Publishing.
18. Li, X., L. Zhu, C.-W. Fu, *et al.* 2020. CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. *IEEE Trans. Med. Imaging* **39:** 1483–1493.
19. Pfister, M., K. Schützenberger, U. Pfeiffenberger, *et al.* 2019. Automated segmentation of dermal fillers in OCT images of mice using convolutional neural networks. *Biomed. Opt. Express* **10:** 1315–1328.
20. Jia, Y., O. Tan, J. Tokayer, *et al.* 2012. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. *Opt. Express* **20:** 4710–4725.
21. He, K., X. Zhang, S. Ren, *et al.* 2016. Identity mappings in deep residual networks. In ECCV 2016: Computer Vision-ECCV 2016. B. Leibe, J. Matas, N. Sebe & M. Welling, Eds.: 630–645. Springer International Publishing AG.
22. Ronneberger, O., P. Fischer & T. Brox. 2015. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015.* N. Navab, J. Hornegger, W.M. Wells *et al.*, Eds.: 234–241. Cham: Springer International Publishing.
23. Chollet, F. 2015. Keras. Accessed January 29, 2021. https://keras.io.
24. Abadi, M., A. Agarwal, P. Barham, *et al.* 2015. TensorFlow: large-scale machine learning on heterogeneous systems. In OSDI'16: Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation. 265–283.
25. Simonyan, K., A. Vedaldi & A. Zisserman. 2014. Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv:1312.6034 [cs]*; https://iclr.cc/archive/2014/workshop-proceedings/.
26. Erhan, D., Y. Bengio, A. Courville, *et al.* 2009. *Visualizing Higher-Layer Features of a Deep Network*. University of Montreal.
27. Baehrens, D., T. Schroeter, S. Harmeling, *et al.* 2010. How to explain individual classification decisions. *J. Mach. Learn. Res.* **11:** 1803–1831.
28. Adebayo, J., J. Gilmer, M. Muelly, *et al.* 2018. Sanity checks for saliency maps. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018).* S. Bengio, H. Wallach, H. Larochelle, et al., Eds.: 9525–9536. Montreal, Canada.
29. Smilkov, D., N. Thorat, B. Kim, *et al.* 2017. SmoothGrad: removing noise by adding noise. *arXiv:1706.03825 [cs, stat]*; https://icmlviz.github.io/papers/ and https://dblp.org/rec/journals/corr/SmilkovTKVW17.html.
30. Chu, Z., J. Lin, C. Gao, *et al.* 2016. Quantitative assessment of the retinal microvasculature using optical coherence tomography angiography. *J. Biomed. Opt.* **21.** 066008.
31. Sato, Y., S. Nakajima, N. Shiraga, *et al.* 1998. Three-dimensional multi-scale line filter for segmentation and

visualization of curvilinear structures in medical images. *Med. Image Anal.* **2:** 143–168.

32. Pizer, S.M., E.P. Amburn, J.D. Austin, *et al.* 1987. Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process.* **39:** 355–368.

33. Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9:** 62–66.

34. Zhang, T.Y. & C.Y. Suen. 1984. A fast parallel algorithm for thinning digital patterns. *Commun. ACM* **27:** 236–239.

35. van der, W.S., J.L. Schönberger, J. Nunez-Iglesias, *et al.* 2014. Scikit-image: image processing in Python. *PeerJ* **2.** e453.

36. Nunez-Iglesias, J., A.J. Blanch, O. Looker, *et al.* 2018. A new Python library to analyse skeleton images confirms malaria parasite remodelling of the red blood cell membrane skeleton. *PeerJ* **6.** e4312.

37. Tellechea, A., A. Kafanas, E.C. Leal, *et al.* 2013. Increased skin inflammation and blood vessel density in human and experimental diabetes. *Int. J. Low. Extrem. Wounds* **12:** 4–11.

38. Tam, J., K.P. Dhamdhere, P. Tiruveedhula, *et al.* 2011. Disruption of the retinal parafoveal capillary network in type 2 diabetes before the onset of diabetic retinopathy. *Invest. Ophthalmol. Vis. Sci.* **52:** 9257–9266.

39. Couturier, A., V. Mané, S. Bonnin, *et al.* 2015. Capillary plexus anomalies in diabetic retinopathy on optical coherence tomography angiography. *Retina* **35:** 2384–2391.

40. Jörneskog, G., K. Brismar & B. Fagrell. 1995. Skin capillary circulation is more impaired in the toes of diabetic than nondiabetic patients with peripheral vascular disease. *Diabet. Med.* **12:** 36–41.

41. Fiordaliso, F., G. Clerici, S. Maggioni, *et al.* 2016. Prospective study on microangiopathy in type 2 diabetic foot ulcer. *Diabetologia* **59:** 1542–1548.