Machine learning analysis of whole mouse brain vasculature

Mihail Ivilinov Todorov[1,2,3,10], Johannes Christian Paetzold[4,5,6,10], Oliver Schoppe[4,5], Giles Tetteh[4], Suprosanna Shit[4,5,6], Velizar Efremov[4,7], Katalin Todorov-Völgyi[2], Marco Düring[2,8], Martin Dichgans[2,8,9], Marie Piraud[4], Bjoern Menze [4,5,6,11✉] and Ali Ertürk [1,2,8,11✉]

Tissue clearing methods enable the imaging of biological specimens without sectioning. However, reliable and scalable analysis of large imaging datasets in three dimensions remains a challenge. Here we developed a deep learning-based framework to quantify and analyze brain vasculature, named Vessel Segmentation & Analysis Pipeline (VesSAP). Our pipeline uses a convolutional neural network (CNN) with a transfer learning approach for segmentation and achieves human-level accuracy. By using VesSAP, we analyzed the vascular features of whole C57BL/6J, CD1 and BALB/c mouse brains at the micrometer scale after registering them to the Allen mouse brain atlas. We report evidence of secondary intracranial collateral vascularization in CD1 mice and find reduced vascularization of the brainstem in comparison to the cerebrum. Thus, VesSAP enables unbiased and scalable quantifications of the angioarchitecture of cleared mouse brains and yields biological insights into the vascular function of the brain.

Changes in cerebrovascular structures are key indicators for a large number of diseases affecting the brain. Primary angiopathies, vascular risk factors (for example, diabetes), traumatic brain injury, vascular occlusion and stroke all affect the function of the brain's vascular network[1–3]. The hallmarks of Alzheimer's disease, including tauopathy and amyloidopathy, can also lead to aberrant remodeling of blood vessels[1,4], allowing capillary rarefaction to be used as a marker for vascular damages[5]. Therefore, quantitative analysis of the entire brain vasculature is pivotal to developing a better understanding of brain function in physiological and pathological states. However, quantifying micrometer-scale changes in the cerebrovascular network of the brain has been difficult for two main reasons.

First, labeling and imaging of the complete mouse brain vasculature down to the smallest blood vessels has not yet been achieved. Magnetic resonance imaging (MRI), micro-computed tomography (micro-CT) and optical coherence tomography do not have sufficient resolution to capture capillaries in bulk tissue[6–8]. Fluorescent microscopy provides higher resolution, but can typically only be applied to tissue sections up to 200 μm in thickness[9]. Recent advances in tissue clearing could overcome this problem[10], but so far there has been no systematic description of all vessels of all sizes in an entire brain in three dimensions (3D).

The second challenge relates to the automated analysis of large 3D imaging datasets with substantial variance in signal intensity and signal-to-noise ratio (SNR) at different depths. Simple intensity- and shape-based filtering approaches such as Frangi's vesselness filters and more advanced image processing methods with local spatial adaptation cannot reliably differentiate vessels from background in whole-brain scans[11,12]. Finally, imaging of the complete vascular network of the brain at capillary resolution results in datasets of terabyte size. Established image processing methods do not scale well to terabyte-sized image volumes, as they do not generalize well to large images, and require intensive manual fine-tuning[13–15].

Here we present VesSAP (Vessel Segmentation & Analysis Pipeline), a deep learning-based method for automated analysis of the entire mouse brain vasculature, overcoming the above limitations. VesSAP encompasses three major steps: (1) staining, clearing and imaging of the mouse brain vasculature down to the capillary level with two different dyes: wheat germ agglutinin (WGA) and Evans blue (EB); (2) automatic segmentation and tracing of the whole-brain vasculature data via CNNs; and (3) extraction of vascular features for hundreds of brain regions after registration of the data to the Allen brain atlas (Fig. 1). Our deep learning-based approach for network extraction in cleared tissue is robust, despite variations in signal intensities and structures, outperforms previous filter-based methods and reaches the quality of segmentation achieved by human annotators. We applied VesSAP to the three commonly used mouse strains C57BL/6J, CD1 and BALB/c.

Results

Vascular staining, DISCO clearing and imaging. To reliably stain the entire vasculature, we used WGA and EB dyes, which can be visualized in different fluorescence channels. We injected EB dye into live mice 12 h before WGA perfusion, allowing its long-term circulation to mark vessels under physiological conditions[16], while we perfused mice with WGA during fixation. We then performed
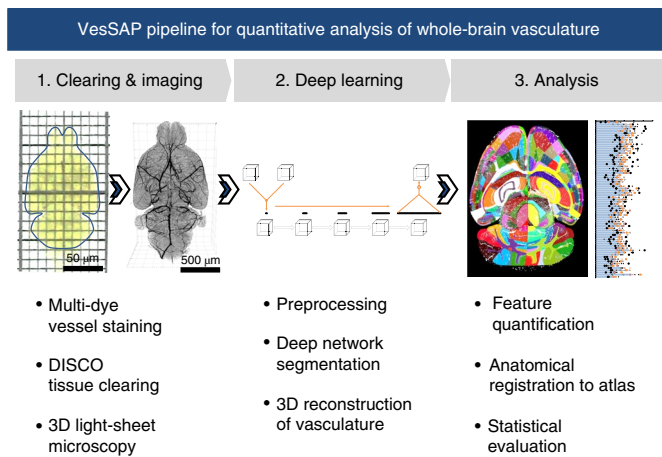
Fig. 1 | Summary of the VesSAP pipeline. The method consists of three modular steps: (1) multi-dye vessel staining and DISCO tissue clearing for high imaging quality using 3D light-sheet microscopy; (2) deep learning-based segmentation of blood vessels with 3D reconstruction; and (3) anatomical feature extraction and mapping of the entire vasculature to the Allen adult mouse brain atlas for statistical analysis.

3DISCO clearing[17] and light-sheet microscopy imaging of whole mouse brains (Fig. 2a–c and Supplementary Figs. 1 and 2). WGA highlighted microvessels, and EB predominantly stained major blood vessels, such as the middle cerebral artery and the circle of Willis (Fig. 2d–i and Supplementary Fig. 3). Merging the signals from the two dyes yielded more complete staining of the vasculature than relying on individual dyes alone (Fig. 2c,f and Supplementary Video 1). Staining with the two dyes was congruent in midsized vessels, with signals originating from the vessel wall layer (Fig. 2j–l and Supplementary Fig. 3a–c). When using WGA, we reached a higher SNR for microvessels than for bigger vessels. With EB, the SNR for small capillaries was lower but larger vessels reached a high SNR (Supplementary Fig. 4). Integrating the information from the two channels allowed acquisition of the entire vasculature and resulted in optimized SNR. We also compared the fluorescence signal quality of the WGA staining (targeting the complete endothelial glycocalyx lining[18]) to signal for a conventional vessel-specific antibody (anti-CD31, targeting endothelial cell–cell adhesion[19]) and found that WGA produced higher SNR for blood vessels in general (Supplementary Fig. 5).

Segmentation of volumetric images. To enable extraction of quantitative features of the vascular structure, vessels in acquired brain scans need to be segmented in 3D. Motivated by deep learning-based approaches in biomedical image data analysis[20–28], we used a five-layer CNN (Fig. 3a) to exploit the complementary signals of the two dyes to derive complete segmentation of the entire brain vasculature.

In the first step, the two input channels (WGA and EB) were concatenated. This yielded a matrix in which each voxel was characterized by two features. Then, each convolutional step integrated the information from a voxel's 3D neighborhood. We used full 3D convolutions[20] without further down- or upsampling and fewer trainable parameters than, for example, 3D U-Net and V-Net[29,30] to achieve high inference speeds. After the fourth convolution, the information from 50 features per voxel was combined with a convolutional layer with a kernel size of one and sigmoidal activation to estimate the likelihood that a given voxel represented a vessel. Subsequent binarization yielded the final segmentation. In both training and testing, the images were processed in subvolumes of $50 \times 100 \times 100$ pixels.
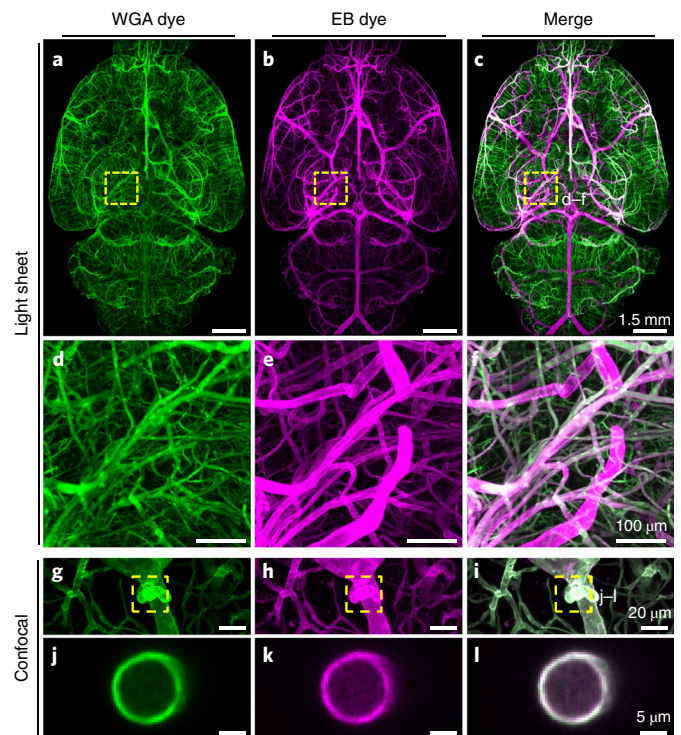

Fig. 2 | Enhancement of vascular staining using two complementary dyes. a–c, Maximum-intensity projections of automatically reconstructed tiling scans of WGA (**a**) and EB (**b**) signal in the same sample and the merged view (**c**). **d–f,** Magnified view of the boxed region in **c**. **g–l,** Confocal images of WGA- and EB-stained vessels and vascular wall (**g–i,** maximum-intensity projections at 112 μm; **j–l,** single slices of 1 μm corresponding to the boxed region in **i**). The experiment was performed on nine different mice with similar results.

Deep neural networks often require large amounts of annotated data or many iterations of training. Here we circumvented this requirement with a transfer learning approach[31]. In short, we first pretrained the network on a large, synthetically generated vessel-like dataset (Supplementary Fig. 6)[32] and then refined it on a small number of manually annotated parts of real brain vessel scans. This approach reduced the training iterations on manually annotated training data.

To assess the quality of the segmentation, we compared the VesSAP CNN predictions to manually labeled ground truth and the predictions from alternative computational approaches (Table 1). We report voxel-wise segmentation metrics, namely, accuracy, F1 score[33], Jaccard coefficient and cl-F1, which weights the centerlines and volumes of the vessels (detailed in the Methods). In comparison to the ground truth, our network achieved an accuracy of 0.94 ± 0.01 and an F1 score of 0.84 ± 0.05 (for additional scores, see Table 1; all values are given as the mean $\pm$ s.d.). As controls, we implemented alternative state-of-the-art deep learning and classical methods. Our network outperformed classical Frangi filters[11] (accuracy, 0.85 ± 0.03; F1 score, 0.47 ± 0.18), as well as recent methods based on local spatial context via Markov random fields[13,34] (accuracy, 0.85 ± 0.03; F1 score, 0.48 ± 0.04). VesSAP achieved similar performance in comparison to 3D U-Net and V-Net architectures, which require substantially more trainable parameters (3D U-Net: accuracy, 0.95 ± 0.01; F1 score, 0.85 ± 0.03; V-Net: accuracy, 0.95 ± 0.02; F1 score, 0.86 ± 0.07; no statistical difference in comparison to the VesSAP CNN: two-sided t test, all $P > 0.3$). However, the VesSAP CNN substantially outperformed the other architectures in terms of speed, being ~20 and ~50 times faster in the feedforward path than
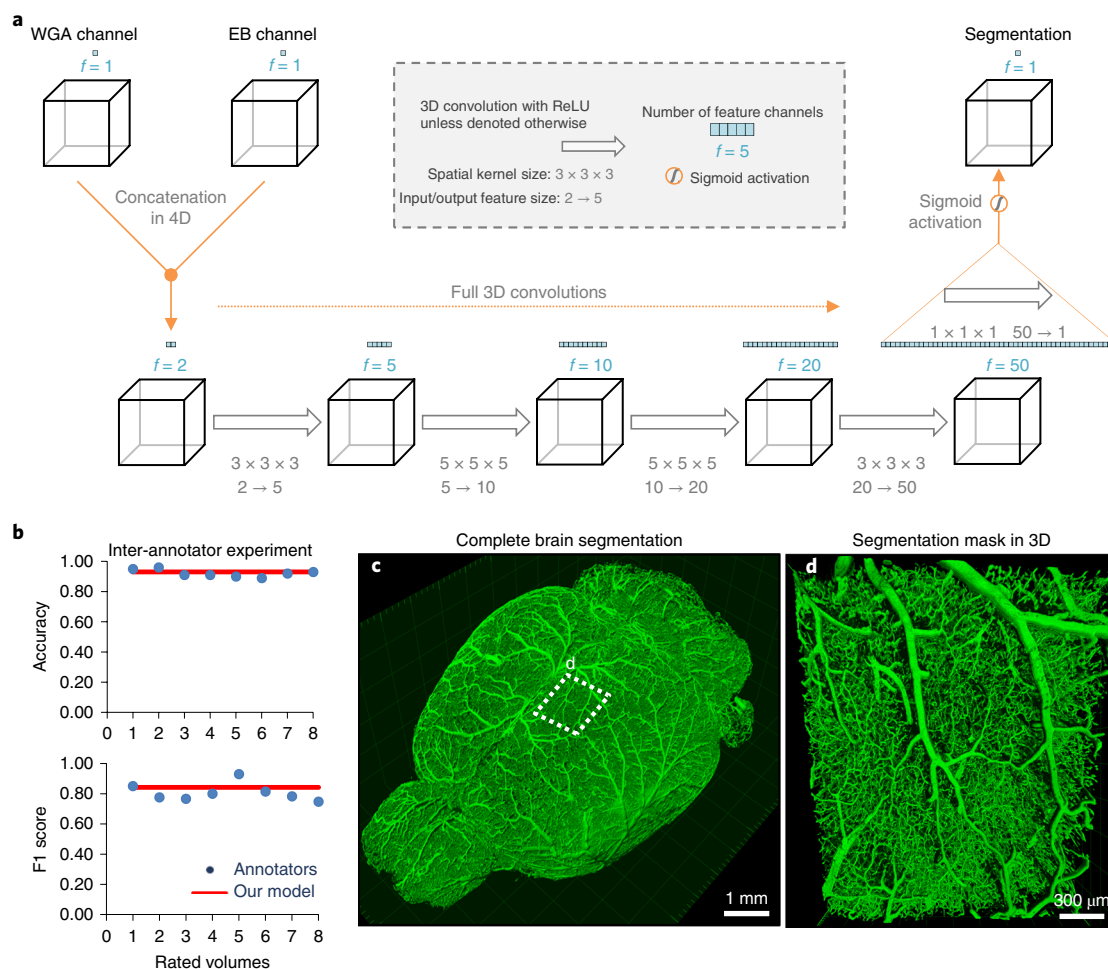
Fig. 3 | Deep learning architecture of VesSAP and performance on vessel segmentation. a, The 3D VesSAP network architecture consisting of five convolutional layers and sigmoid activation for the last layer, including the kernel size and feature size for the input/output. ReLU, rectified linear units. **b**, Accuracy and F1 score for the inter-annotator experiment (blue) as compared to VesSAP (red). **c**, 3D rendering of full brain segmentation from a CD1 mouse. **d**, 3D rendering of the small volume boxed in **c**. The experiment was performed on nine different mice with similar results.

Table 1 | Evaluation metrics of the different segmentation approaches for 75 volumes of 100 × 100 × 50 pixels

Segmentation model	cl-F1	Accuracy	F1 score	Jaccard	Parameters	Speed
VesSAP CNN	**0.93 ± 0.02***	0.94 ± 0.01	0.84 ± 0.05	0.84 ± 0.04	**0.0587 M***	**1.19 s***
VesSAP CNN, trained from scratch	**0.93 ± 0.02**	0.94 ± 0.01	**0.85 ± 0.04***	0.85 ± 0.04	**0.0587 M***	**1.19 s***
VesSAP CNN, synthetic training data	0.87 ± 0.02	0.90 ± 0.05	0.72 ± 0.07	0.70 ± 0.05	**0.0587 M***	**1.19 s***
3D U-Net	**0.93 ± 0.02**	**0.95 ± 0.01***	**0.85 ± 0.03***	0.85 ± 0.03	178.4537 M	61.22 s
V-Net	**0.94 ± 0.02***	**0.95 ± 0.02***	**0.86 ± 0.07***	**0.86 ± 0.07***	88.8556 M	26.87 s
Frangi vesselness	0.84 ± 0.03	0.85 ± 0.03	0.47 ± 0.19	–	–	117.00 s
Markov random field	0.86 ± 0.02	0.85 ± 0.03	0.48 ± 0.04	–	–	24.31 s

All values are given as the mean ± s.d. The best performing algorithms are in bold and highlighted with an asterisk; algorithms whose performance did not differ more than 2% from the best performing algorithms are in bold. The number of trainable parameters for deep learning architectures is given in millions (M).

V-Net and 3D U-Net, respectively. This is particularly important for our large datasets (hundreds of gigabytes). For example, the VesSAP CNN segmented a single brain in 4 h, whereas V-Net and 3D U-Net required 3.3 d and 8 d, respectively. The superior speed of the VesSAP CNN is due to the substantially fewer trainable parameters in its architecture (for example, our implementation of 3D U-Net had ~178 million parameters, whereas the VesSAP CNN had ~0.059

million parameters) (Table 1). Next, we compared the segmentation accuracy of our network to the accuracy of human annotations. A total of four human experts independently annotated two volumes. We found that the inter-annotator accuracy and F1 scores of the experts were comparable to those from the predicted segmentation of our network (human annotators: accuracy, 0.92 ± 0.02; F1 score, 0.81 ± 0.06; Fig. 3b). Notably, we extrapolate that human annotators
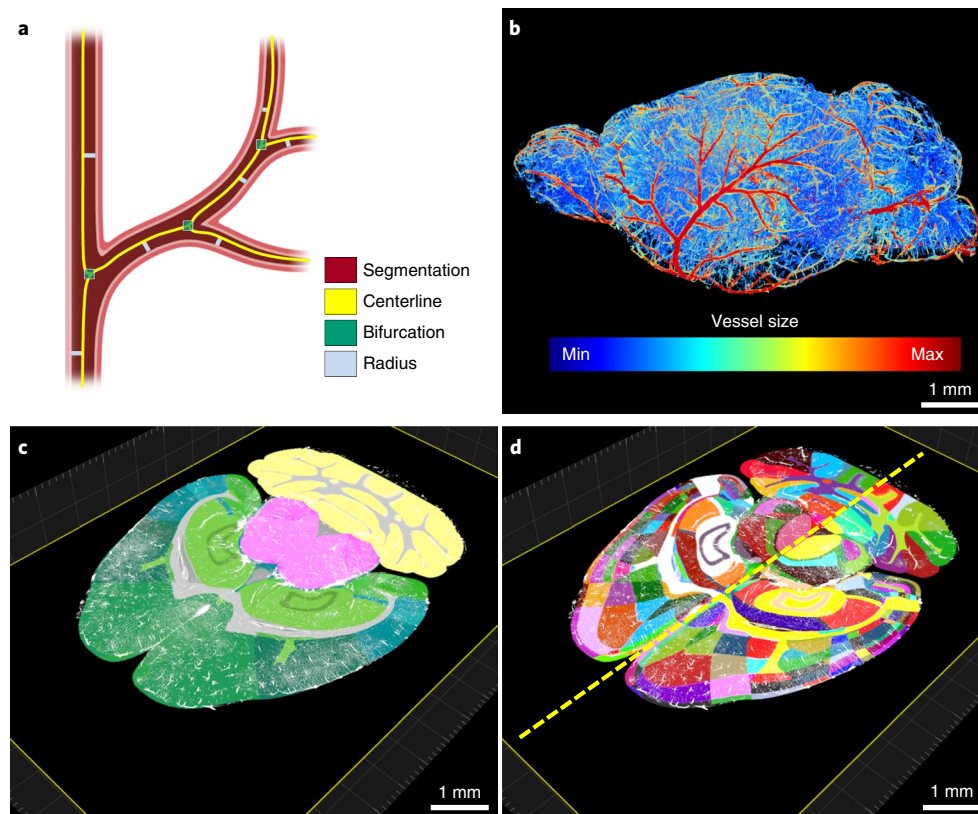
Fig. 4 | Pipeline showing the feature extraction and registration process. a, Representation of the features extracted from vessels. **b**, Radius illustration of the vasculature in a CD1 mouse brain. **c,d**, Vascular segmentation results overlaid on the hierarchically (**c**) and randomly (**d**) color-coded atlas to reveal all annotated regions available, including hemispheric difference (dashed line in **d**). The experiment was performed on nine different mice with similar results.

would need more than a year to process a whole brain instead of the 4 h required by our approach. Moreover, we observed differences in the human segmentations due to annotator bias. Thus, the VesSAP CNN can segment the complete brain vasculature consistently at human-level accuracy with a substantially higher speed than currently available methods, enabling high throughput for large-scale analysis.

We show an example of the vasculature from a brain segmented by VesSAP in 3D (Fig. 3c and Supplementary Videos 2 and 3). Zooming in on a smaller patch revealed that the connectivity of the vascular network was fully maintained (Fig. 3d and Supplementary Video 2). Comparing single slices of the imaging data with the predicted segmentation showed that vessels were accurately segmented regardless of absolute illumination or vessel diameter (Supplementary Fig. 7).

Feature extraction and atlas registration. Vessel lengths and radii and the number of bifurcation points are commonly used to describe the angioarchitecture[2]. Hence, we used our segmentation to quantify these features as distinct parameters to characterize the mouse brain vasculature (Fig. 4a and Supplementary Video 4). We evaluated the local vessel length (length normalized to the size of the brain region of interest), local bifurcation density (sum of the occurrences normalized to the size of the brain region of interest) and local vessel radius (average radius along the full length) of blood vessels in different brain regions.

We report the vascular features in three ways to enable comparison with various previous studies that differed in the measures used (Supplementary Fig. 8). More specifically, first, we provide the count of segmented voxels as compared to total voxels within a specific brain region (voxel space). Second, we provide the

measurements by calculating the voxel size of our imaging system and accounting for the Euclidean length (microscopic space). Third, we corrected the microscopic measurements to account for tissue shrinkage caused by the clearing process (anatomical space)[35,36] (Supplementary Tables 2–10). We calculated this shrinkage rate by measuring the same mouse brain volume with MRI before clearing.

Here we use the anatomical space to report our specific biological findings, as it is closest to the physiological state. For the average blood vessel length of the whole brain, we found a value of 545.74 ± 94 mm per mm^3 (mean ± s.d.). Because our method quantifies brain regions separately, we could compare our results to the literature, which mostly reports either quantifications for specific brain regions or extrapolations to the whole brain from regional quantifications. For example, a vascular length of 922 ± 176 mm per mm^3 (mean ± s.d.) was previously reported for cortical regions (size of $508 \times 508 \times 1{,}500$ µm^3)[10]. We found a similar vessel length for the same region in the mouse cortex (C57BL/6J mice: 913 ± 110 mm per mm^3), substantiating the accuracy of our method. We performed additional comparisons to other reports (Supplementary Table 11). Moreover, we compared the measurements acquired with our algorithms to manually labeled ground truth data and found deviations of 8.21% for centerlines, 13.18% for the number of bifurcation points and 16.33% for the average radius. These deviations were substantially lower than the average deviation among human annotators (Methods).

We quantified and visualized vessel radius along the entire vascular network (Fig. 4b). After extracting vascular features for the whole brain with VesSAP, we registered the volume to the Allen brain atlas (Supplementary Videos 5 and 6). This allowed us to map the segmented vasculature and corresponding features topographically to distinct anatomical brain regions (Fig. 4c). Each anatomical

Fig. 5 | Anatomical properties of the neurovasculature in adult mouse brain mapped to the Allen brain atlas clusters. a–c, Representations of the local vessel length (**a**), density of bifurcations (**b**) and average radius (**c**) in each of the 71 main anatomical clusters of the Allen brain atlas. Open, black and orange circles denote measurements in the CD1, C57BL/6J and BALB/c strains, respectively; each circle represents a single mouse. Data are given as the mean ± s.e.m.; *n* = 3 mice per strain. **d,** Local distribution of large, intermediate and microvessels in the same anatomical clusters. Abbreviations are defined in Supplementary Table 1.

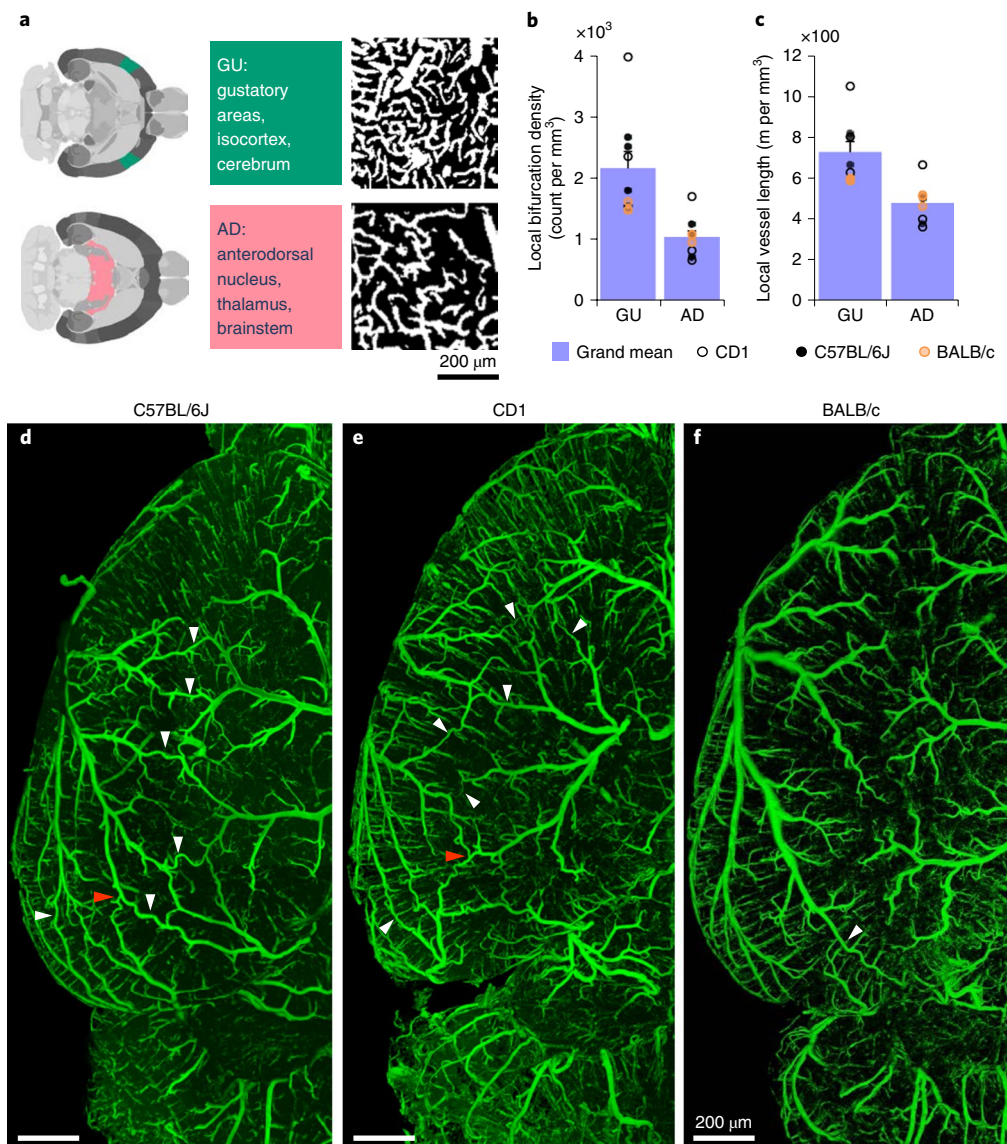**Fig. 6 | Exemplary quantitative analysis enabled by VesSAP. a**, Respective locations of the anterodorsal nucleus (AD) and gustatory areas (GU) in the mouse brain (left) and maximum-intensity projections of representative volumes from segmentation of these areas ($600 \times 600 \times 33 \mu m^3$) (right). **b,c**, Quantification of the bifurcation density (**b**) and local vessel length (**c**) for the anterodorsal nucleus and gustatory area clusters. CD1 mice are shown by open circles, BALB/C mice by orange circles and C57BL/6J mice by black circles. Values are the mean ± s.e.m.; $n = 3$ mice per strain. **d–f**, Images of the vasculature in representative C57BL/6J (**d**), CD1 (**e**) and BALB/c (**f**) mice, where white arrowheads indicate anastomoses between major arteries. Direct vascular connections between the medial cerebral artery, the anterior cerebral artery and the posterior cerebral artery are indicated by red arrowheads. The experiment was performed three times with similar results.

region could be further divided into subregions, yielding a total of 1,238 anatomical structures (619 per hemisphere) for the entire mouse brain (Fig. 4d). This allowed analysis of each denoted brain region and grouping of regions into clusters such as left versus right hemisphere, gray versus white matter, or any hierarchical cluster of the Allen brain atlas ontology. For our subsequent statistical feature analysis, we grouped the labeled structures according to the 71 main anatomical clusters of the current Allen brain atlas ontology. We thus provide the whole mouse brain vascular map with extracted vessel lengths, bifurcation points and radii down to the capillary level.

VesSAP provides a reference map of the whole brain vasculature in mice. By studying whole brain vasculature in the C57BL/6J, CD1 and BALB/c strains ($n = 3$ mice for each strain), we found that the local vessel length and local bifurcation density differed in the same

brain over different regions, while they were highly correlated among different mice for the same regions (Fig. 5a,b). Furthermore, the local bifurcation density was highly correlated with the local vessel length in most brain regions (Supplementary Fig. 9), and the average vessel radius was evenly distributed in different regions of the same brain (Fig. 5c). In addition, the extracted features showed no statistical difference (by Cohen's d; Supplementary Table 12) for the same anatomical cluster across the strains (Supplementary Fig. 9). Finally, microvessels made up the overwhelming majority of the total vascular composition in all brain regions (Fig. 5d). We visually inspected exemplary brain regions to validate the output of VesSAP. Both VesSAP and visual inspection revealed that the gustatory areas had a higher vascular length per volume than the anterodorsal nucleus (Fig. 6a–c). Visual inspection also suggested that the number of capillaries was the primary reason for regional feature variations within the same brain.

Finally, VesSAP offered insights into the neurovascular structure of the different mouse strains in our study. There were direct intracranial vascular anastomoses in the C57BL/6J, CD1 and BALB/c strains (white arrowheads in Fig. 6d–f). The anterior cerebral artery, middle cerebral artery and posterior cerebral artery were connected at the dorsal visual cortex in CD1 mice (red arrowheads in Fig. 6d,e) unlike in the BALB/c strain[33] (Fig. 6f).

Discussion

VesSAP can generate reference maps of the adult mouse brain vasculature, which can potentially be used to model synthetic cerebrovascular networks[37]. In addition to the metrics we obtain to describe the vasculature, advanced metrics, for example, Strahler values, network connectivity and bifurcation angles, can be extracted by using the data generated by VesSAP. Furthermore, the centerlines and bifurcation points can be interpreted as the edges and nodes for building a full vascular network graph, offering a means for studying local and global properties of the cerebrovascular network in the future.

The VesSAP workflow relies on staining of blood vessels by two different dyes. WGA binds to the glycocalyx of the endothelial lining of blood vessels[38] but may miss some segments of large vessels[18]. EB is a dye with a high affinity for serum albumin[35,36,39]; thus, it remains in the large vessels after a short perfusion protocol. In addition, EB labeling is not affected by subsequent DISCO clearing.

Vessels have long and thin tubular shapes. In our images, the radii of capillaries (about 3 μm) are in the range of our voxel size. Therefore, segmentation that yields the correct diameter down to single-pixel resolution poses a challenge, as we observed a 16% deviation for the radius. This subpixel deviation did not pose a problem for segmenting the whole vasculature network and extracting features because the vascular network can be defined by its centerlines and bifurcations.

The described segmentation concept is based on a transfer learning approach, where we pretrained the CNN and refined it on a small labeled dataset of 11% of the synthetic dataset and only 0.02% of one cleared brain. We consider this to be a major advantage in comparison to training from scratch. Thus, our CNN might generalize well to different types of imaging data (such as micro-CT angiography) or other curvilinear structures (for example, neurons), as only a small labeled dataset is needed to adjust our pretrained network.

On the basis of our vascular reference map, unknown vascular properties can be discovered and biological models can be confirmed. VesSAP showed a high number of collaterals in albino CD1 mice. Such collaterals between large vessels can substantially alter the outcome of ischemic stroke lesions: blood-deprived brain regions arising from occlusion of a large vessel can be compensated by blood supply from the collateral extensions of other large vessels[33,40]. Therefore, our VesSAP method can lead to the discovery of previously unknown anatomical details that could be functionally relevant.

In conclusion, VesSAP is a scalable, modular and automated machine learning-based method to analyze complex imaging data from cleared mouse brains. We foresee that our method will accelerate the applications of tissue clearing, in particular for studies assessing brain vasculature.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgments, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-020-0792-1.

References

1. Bennett, R. E. et al. Tau induces blood vessel abnormalities and angiogenesis-related gene expression in P301L transgenic mice and human Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **115**, E1289–E1298 (2018).
2. Obenaus, A. et al. Traumatic brain injury results in acute rarefication of the vascular network. *Sci. Rep.* **7**, 239 (2017).
3. Völgyi, K. et al. Chronic cerebral hypoperfusion induced synaptic proteome changes in the rat cerebral cortex. *Mol. Neurobiol.* **55**, 4253–4266 (2018).
4. Klohs, J. et al. Contrast-enhanced magnetic resonance microangiography reveals remodeling of the cerebral microvasculature in transgenic ArcAbeta mice. *J. Neurosci.* **32**, 1705–1713 (2012).
5. Edwards-Richards, A. et al. Capillary rarefaction: an early marker of microvascular disease in young hemodialysis patients. *Clin. Kidney J.* **7**, 569–574 (2014).
6. Calabrese, E., Badea, A., Cofer, G., Qi, Y. & Johnson, G. A. A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data. *Cereb. Cortex* **25**, 4628–4637 (2015).
7. Dyer, E. L. et al. Quantifying mesoscale neuroanatomy using X-ray microtomography. *eNeuro* **4**, ENEURO.0195-17.2017 (2017).
8. Li, T., Liu, C. J. & Akkin, T. Contrast-enhanced serial optical coherence scanner with deep learning network reveals vasculature and white matter organization of mouse brain. *Neurophotonics* **6**, 035004 (2019).
9. Tsai, P. S. et al. Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels. *J. Neurosci.* **29**, 14553–14570 (2009).
10. Lugo-Hernandez, E. et al. 3D visualization and quantification of microvessels in the whole ischemic mouse brain using solvent-based clearing and light sheet microscopy. *J. Cereb. Blood Flow Metab.* **37**, 3355–3367 (2017).
11. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 130–137 (Springer, 1998).
12. Sato, Y. et al. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med. Image Anal.* **2**, 143–168 (1998).
13. Di Giovanna, A. P. et al. Whole-brain vasculature reconstruction at the single capillary level. *Sci. Rep.* **8**, 12573 (2018).
14. Xiong, B. et al. Precise cerebral vascular atlas in stereotaxic coordinates of whole mouse brain. *Front. Neuroanat.* **11**, 128 (2017).
15. Clark, T. A. et al. Artery targeted photothrombosis widens the vascular penumbra, instigates peri-infarct neovascularization and models forelimb impairments. *Sci. Rep.* **9**, 2323 (2019).
16. Yao, J., Maslov, K., Hu, S. & Wang, L. V. Evans blue dye—enhanced capillary-resolution photoacoustic microscopy in vivo. *J. Biomed. Opt.* **14**, 054049 (2009).
17. Ertürk, A. et al. Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nat. Protoc.* **7**, 1983–1995 (2012).
18. Reitsma, S. et al. Endothelial glycocalyx structure in the intact carotid artery: a two-photon laser scanning microscopy study. *J. Vasc. Res.* **48**, 297–306 (2011).
19. Schimmenti, L. A., Yan, H. C., Madri, J. A. & Albelda, S. M. Platelet endothelial cell adhesion molecule, PECAM-1, modulates cell migration. *J. Cell. Physiol.* **153**, 417–428 (1992).
20. Tetteh, G. et al. DeepVesselNet: vessel segmentation, centerline prediction, and bifurcation detection in 3-D angiographic volumes. Preprint at https://arxiv.org/abs/1803.09340 (2018).
21. Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).
22. Wang, H. et al. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods* **16**, 103–110 (2019).
23. Falk, T. et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).
24. Caicedo, J. C. et al. Data-analysis strategies for image-based cell profiling. *Nat. Methods* **14**, 849–863 (2017).
25. Dorkenwald, S. et al. Automated synaptic connectivity inference for volume electron microscopy. *Nat. Methods* **14**, 435–442 (2017).
26. Liu, S., Zhang, D., Song, Y., Peng, H. & Cai, W. in *Machine Learning in Medical Imaging* (eds., Wang, Q. et al.) 185–193 (Springer International Publishing, 2017).
27. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 3431–3440 (IEEE, 2015).
28. Livne, M. et al. A U-Net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease. *Front. Neurosci.* **13**, 97 (2019).
29. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T. & Ronneberger, O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. in *Medical Image Computing and Computer-Assisted Intervention* 424–432 (2016).
30. Milletari, F., Navab, N. & Ahmadi, S.A. V-Net: fully convolutional neural networks for volumetric medical image segmentation. in *2016 Fourth International Conference on 3D Vision* 565–571 (IEEE, 2016).

31. Litjens, G. et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88 (2017).
32. Schneider, M., Reichold, J., Weber, B., Szekely, G. & Hirsch, S. Tissue metabolism driven arterial tree generation. *Med. Image Anal.* **16**, 1397–1414 (2012).
33. Chalothorn, D., Clayton, J. A., Zhang, H., Pomp, D. & Faber, J. E. Collateral density, remodeling, and VEGF-A expression differ widely between mouse strains. *Physiol. Genomics* **30**, 179–191 (2007).
34. Li, S. Z. In *Computer Vision—ECCV* (ed., Eklundh, J.-O.) 361–370 (Springer, 1994).
35. Pan, C. et al. Shrinkage-mediated imaging of entire organs and organisms using uDISCO. *Nat. Methods* **13**, 859–867 (2016).
36. Cai, R. et al. Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull–meninges connections. *Nat. Neurosci.* **22**, 317–327 (2019).
37. Menti, E., Bonaldi, L., Ballerini, L., Ruggeri, A. & Trucco, E. In *International Workshop on Simulation and Synthesis in Medical Imaging* 167–176 (Springer, Year).
38. Kataoka, H. et al. Fluorescent imaging of endothelial glycocalyx layer with wheat germ agglutinin using intravital microscopy. *Microsc. Res. Tech.* **79**, 31–37 (2016).
39. Steinwall, O. & Klatzo, I. Selective vulnerability of the blood–brain barrier in chemically induced lesions. *J. Neuropathol. Exp. Neurol.* **25**, 542–559 (1966).
40. Zhang, H., Prabhakar, P., Sealock, R. & Faber, J. E. Wide genetic variation in the native pial collateral circulation is a major determinant of variation in severity of stroke. *J. Cereb. Blood Flow Metab.* **30**, 923–934 (2010).

Methods

Tissue preparation. Animal experiments were conducted according to institutional guidelines (Klinikum der Universität München/Ludwig Maximilian University of Munich), after approval of the ethical review board of the government of Upper Bavaria (Regierung von Oberbayern, Munich, Germany), and in accordance with European directive 2010/63/EU for animal research. Animals were housed under a 12-h light/12-h dark cycle. For this study, we injected 150 µl (2% (vol/vol) in saline) EB dye (Sigma-Aldrich, E2129) intraperitoneally into 3-month-old male mice of the C57BL/6J, CD1 and BALB/c strains (Charles River, strain codes 027, 482 and 028, respectively; $n = 3$ mice per strain). Twelve hours after injection of EB dye, we anesthetized the animals with a combination of midazolam, medetomidine and fentanyl (administered intraperitoneally; 1 ml per 100 g body weight containing 5 mg, 0.5 mg and 0.05 mg per kg body weight, respectively) and opened their chest for transcardial perfusion. Medium with WGA (0.25 mg WGA conjugated to Alexa Fluor 594 dye (Thermo Fisher Scientific, W11262) in 150 µl PBS, pH 7.2) was supplied by peristaltic pump set to deliver the medium at a rate of 8 ml min⁻¹, along with 15 ml of 1× PBS and 15 ml of 4% paraformaldehyde. This short perfusion protocol was established on the basis of preliminary experiments, where both WGA and EB staining were partially washed out (data not shown), with the goal of delivering fixative to brain tissue via the vessels to achieve a homogenous preservation effect[41].

After perfusion, brains were extracted from the neurocranium while severing some of the segments of the circle of Willis, which is an inevitable component of most retrieval processes aside from corrosion cast techniques. Next, the samples were incubated in 3DISCO clearing solutions as described[17]. Briefly, we immersed them in a gradient of tetrahydrofuran (Sigma-Aldrich, 186562): 50%, 70%, 80% and 90% (in distilled water) followed by 100%, at 25 °C for 12 h at each concentration. At this point, we modified the protocol by incubating the samples in *tert*-butanol for 12 h at 35 °C followed by immersion in dichloromethane (Sigma-Aldrich, 270997) for 12 h at room temperature and a final incubation with refractive index-matched BABB solution (benzyl alcohol + benzyl benzoate, 1:2; Sigma-Aldrich, 24122 and W213802), for at least 24 h at room temperature until transparency was achieved. Each incubation step was carried out on a laboratory shaker.

Imaging of cleared samples with light-sheet microscopy. We used a ×4 objective lens (Olympus XLFLUOR 340) equipped with an immersion-corrected dipping cap mounted on a LaVision UltraII microscope coupled to a white-light laser module (NKT SuperK Extreme EXW-12) for imaging. Images were taken with 16-bit depth and at a nominal resolution of 1.625 µm per voxel on the x and y axes. For ×12 imaging, we used a LaVision objective (×12/0.53 NA MI PLAN with an immersion-corrected dipping cap). Brain structures were visualized by Alexa Fluor 594 (using a 580/25-nm excitation filter and a 625/30-nm emission filter) and EB fluorescent dye (using a 640/40-nm excitation filter and a 690/50-nm emission filter) in sequential order. We maximized the SNR for each dye independently to avoid saturation of differently sized vessels when only a single channel was used. We achieved this by independently optimizing the excitation power so that the strongest signal in major vessels did not exceed the dynamic range of the camera. In the z dimension, we took sectional images in 3-µm steps while using left- and right-sided illumination. Our measured resolution was 2.83 µm × 2.83 µm × 4.99 µm for x, y and z, respectively (Supplementary Fig. 2). To reduce defocus, which derives from the Gaussian shape of the beam, we used 12-step sequential shifting of the focal position of the light sheet per plane and side. The thinnest point of the light sheet was 5 µm.

Imaging of cleared samples with confocal microscopy. Additionally, the cleared specimens were imaged with an inverted laser-scanning confocal microscope (Zeiss, LSM 880) for further analysis. Before imaging, samples were mounted by placing them onto the glass surface of a 35-mm glass-bottom Petri dish (MatTek, P35G-0-14-C) and immersed in BABB. A ×40 oil-immersion objective lens was used (Zeiss, ECPlan-NeoFluar ×40/1.30 NA Oil DIC M27, WD = 0.21 mm). Images were acquired with the settings for Alexa Fluor 594 (using excitation at 561 nm and an emission range of 585–733 nm) and EB fluorescent dye (using excitation at 633 nm and an emission range of 638–755 nm) in sequential order.

Magnetic resonance imaging. We used a nanoScan PET/MR device (3 Tesla, Mediso Medical Imaging Systems) equipped with a head coil for murine heads to acquire anatomical scans in the T1 modality.

Reconstruction of the datasets from tiling volumes. We stitched the acquired volumes by using TeraStitcher's automatic global optimization function (v1.10.3). We produced volumetric intensity images of the whole brain while considering each channel separately. To improve alignment to the Allen brain atlas, we downscaled our dataset in the xy plane to achieve pseudouniform voxel spacing matching the z plane.

Deep learning network architecture. We relied on a deep 3D CNN for segmentation of our blood vessel dataset. The network's general architecture consists of five convolutional layers, four with ReLU (rectified linear units) followed by one convolutional layer with sigmoid activation (Fig. 3a). The input layer is designed to take n images as input. In the implemented case, the input to the first layer of the network comprised $n = 2$ images of the same brain, which had been stained differently (Fig. 3a). To specifically account for the general class imbalance (much more tissue background than vessel signal) in our dataset and the potential for high false-positive rates associated with this, we chose the generalized soft-Dice as the loss function to our network. At all levels, we used full 3D convolutional kernels (Fig. 3a).

The network's training is driven by an Adam optimizer with a learning rate of 1×10^{-5} and an exponential decay rate of 0.9 for the first moment and 0.99 for the second moment[42]. A prediction or segmentation with a trained model takes volumetric images of arbitrary size as input and outputs a probabilistic segmentation map of identical size. To deal with volumes of arbitrary size and extension, we processed them in smaller subvolumes of $100 \times 100 \times 50$ pixels in size. The algorithms were implemented by using the Tensorflow framework and KERAS[43]. They were trained and tested on two NVIDIA Quadro P5000 GPUs and on machines with 64 GB and 512 GB of RAM.

Transfer learning. Typically, supervised learning tasks in biomedical imaging are aggravated by the scarce availability of labeled training data. Our transfer learning approach aims to circumvent this problem by pretraining our models on synthetically generated data and refining them on a small set of real images[44]. Specifically, our approach pretrains the VesSAP CNN on 3D volumes of vascular image data, synthetically generated together with the corresponding training labels by using the approach of Schneider and colleagues[45]. The pretraining is carried out on a dataset of 20 volumes of $325 \times 304 \times 600$ pixels in size for 38 epochs. During pretraining, we applied a learning rate of 1×10^{-4}. Afterward, the pretrained model was fine-tuned by retraining on a real microscopic dataset consisting of 11 volumes of $500 \times 500 \times 50$ pixels in size. The image volumes were manually annotated by commissioned experts, including the expert who previously prepared the samples and operated the microscope. The labels were verified and further refined in consensus by two additional human raters. The data we used in this fine-tuning step amounted to 11% of the volume of the synthetic datasets and only 0.02% of the voxel volume of a single whole brain. For the fine-tuning step, we used a learning rate of 1×10^{-5}. The final model was obtained after training on the real dataset for six epochs. This training was substantially shorter than training from scratch, where we trained the same VesSAP CNN architecture for 72 epochs until we reached the best F1 score on the validation set. The labeled dataset consisted of 17 volumes of $500 \times 500 \times 50$ pixels from five mouse brains. Three of these brains were from the CD1 strain, and two were from the C57BL/6J strain. The volumes were chosen from regions throughout the whole brain, to represent the variability in the vascular dataset in terms of both vessel shape and illumination. To ensure independence, volumes for the training set and test/validation set were chosen from independent brains. All datasets included brains from the two strains. Our training dataset consisted of 11 volumes, the validation dataset of 3 volumes and the test dataset of 3 volumes. We cross-tested on our test and validation datasets by rotating these. The volumes were processed during training and inference in 25 small subvolumes of $100 \times 100 \times 50$ pixels.

We observed an average F1 score of 0.84 ± 0.02 (mean ± s.d.), an average accuracy of 0.94 ± 0.01 (mean ± s.d.) and an average Jaccard coefficient of 0.84 ± 0.04 (mean ± s.d.) on our test datasets (Fig. 3b). We tested the statistical significance of differences among the top three learning methods (the VesSAP CNN, V-Net and 3D U-Net) by using paired t tests. We found that the differences in F1 score were not statistically significant (all $P > 0.3$, rejecting the hypothesis of different distributions).

Because the F1 score, accuracy and Jaccard coefficient are all voxel-wise volumetric scores and can fall short in evaluating the connectedness of components, we developed the cl-F1 score. cl-F1 is calculated from the intersection of centerlines and vessel volumes and not from volumes only, as the traditional F1 score is[46]. To determine this score, we first calculated the intersection of the centerline of our prediction with the labeled volume and then calculated the intersection of the labeled volume's centerline with the predicted volume. Next, we treated the first intersection as recall, as it is susceptible to false negatives, and the second intersection as precision, as it is susceptible to false positives, and input this into the traditional F1 score formulation:

$$\text{F1} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{1}$$

We report an average cl-F1 score of 0.93 ± 0.02 (mean ± s.d.) on the test set.

All scores are given as mean and s.d. Our model reached the best model selection point on the validation dataset after six epochs of training.

Comparison to 3D U-Net and V-Net. To compare our proposed architecture to different segmentation architectures, we implemented V-Net and 3D U-Net, both of which use more complex CNNs with substantially more trainable parameters, which further include down- and upsampling. While our experiments showed that 3D U-Net and V-Net reached marginally higher performance scores, the differences were not statistically significant (two-sided t test, $P > 0.3$). The amount

of parameters for these tools makes them 51 and 23 times slower than VesSAP during the inference stage. For segmentation of one of our large whole-brain datasets, this translated to 4 h for VesSAP versus 8 d for 3D U-Net and 3.8 d for V-Net. This difference was also prevalent in the number of trainable parameters. The VesSAP CNN had 0.058 million parameters, whereas 3D U-Net consisted of more than 178 million and V-Net of more than 88 million parameters. Furthermore, the light VesSAP CNN already reached human-level performance. We therefore consider the problem of vessel segmentation as solved by the VesSAP CNN for our data. It should be mentioned that the segmentation network is a modular building block of the overall VesSAP pipeline and can be chosen by each user according to his or her own preferences and, importantly, according to the computational power available.

Preprocessing of segmentation. Preprocessing factors into the overall success of the training and segmentation. The intensity distribution among brains and among brain regions differs substantially. To account for intensity distributions, two preprocessing strategies were applied successively.

1. High-cut filter. In this step, the intensities x above a certain threshold c are set to c; c is defined by a global percentile. Next, they were normalized by $f(x)$.

$$g(x) = \begin{cases} c, & x > c \\ x, & x \le c \end{cases} \quad (2)$$

2. Normalization of intensities. The original intensities were normalized to a range of 0 to 1, where x was the pixel intensity and X was all intensities for the volume.

$$f(x) = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (3)$$

Inter-annotator experiment for segmentation. To compare VesSAP's segmentation to human-level annotations, we implemented an inter-annotator experiment. For this experiment, we determined a gold-standard label (ground truth) for two volumes of $500 \times 500 \times 50$ pixels from a commissioned group of three experts, including the expert who imaged our data and was therefore most familiar with the images. Next, we gave the two volumes to four other experts to label the complete vasculature. The experts spent multiple hours labeling each patch in the ImageJ and ITK-snap environment and were allowed to use their favored approaches to generate what they considered to be the most accurate labeling. Finally, we calculated the accuracy and F1 scores for the different annotators, as compared to the gold standard, and compared them to the scores for our model (Table 1).

Feature extraction. To quantify the anatomy of the mouse brain vasculature, we extracted descriptive features on the basis of our segmentation. First, we calculated the features in voxel space. Later, we registered them to the Allen brain atlas.

As features we extracted the centerlines, the bifurcation points and the radii of the segmented blood vessels. We consider these features to be independent from the elongation of the light-sheet scans and the connectedness of the vessels due to staining, imaging and/or segmentation artifacts.

Our centerline extraction was based on a 3D thinning algorithm[47]. Before extracting the centerlines, we applied two cycles of binary erosion and dilation to remove false-negative pixels within the volume of segmented vessels, as these would induce false centerlines. On the basis of the centerlines, we extracted bifurcation points. A bifurcation was the branching point on a centerline where a larger vessel split into two or more smaller vessels (Fig. 4a). In a network analysis context, bifurcations are meaningful as they represent the nodes of a vascular network[48]. Furthermore, bifurcation points have relevance in a biological context. In neurodegenerative diseases, capillaries are known to degenerate[49], thereby substantially reducing the number of bifurcation points in an affected brain region as compared to healthy brain. Next, we implemented an algorithm to detect bifurcation points. We achieved this by calculating the surrounding pixels for every point on each centerline and determined whether a point was a centerline. The radius of a blood vessel is a key feature to describe vascular networks. The radius yields information about the flow and hierarchy of the vessel network. The proposed algorithm calculates the Euclidean distance transform for every segmented pixel v to the closest background pixel $b_{closest}$. Next, the distance transform matrix is multiplied by the 3D centerline mask, equaling the minimum radius of the vessel around the centerline.

$$d(v, b_{closest}) = \sqrt{\sum_1^3 (v_i - b_{closest,i})^2} \quad (4)$$

Feature quantification. Here we describe in detail how we calculated the features between the three different spaces.

Voxel space to microscopic space. To quantify the length of our vessels in SI units instead of voxels, we calculated their Euclidean length, which depends on the direction of the connection of skeleton pixels (Supplementary Fig. 9). To calculate

the Euclidean length of our centerlines, we carried out a connected component analysis, which transformed each pixel of the skeleton into an element of an undirected weighted graph, where zero weight means no connection and non-zero weights denote the Euclidean distance between two voxels (considering 26 connectivity). Thus, we obtained a large and sparse adjacency matrix. An element-wise summation of such a matrix provides the total Euclidean length of the vascular network along the extracted skeleton.

As measuring connected components is computationally very expensive, we calculated the Euclidean length of the centerlines for 12 representative volumes of $500 \times 500 \times 50$ pixels and divided by the number of skeleton pixels. We calculated an average Euclidean length ε_{Cl} of 1.3234 ± 0.0063 voxels (mean $\pm$ s.d.) per centerline element. This corresponds to a length of $3.9701 \pm 0.0188\,\mu m$ (mean $\pm$ s.d.) in cleared tissue. Because the s.d. of this measurement was low, at less than 0.5% of the length, we applied this correction factor to the whole brain centerline measurements. This correction does not apply to the bifurcation points and our radius statistics, as bifurcations are independent of length and also radius extraction returns a Euclidean distance by default. Depending on the direction of the connection of skeleton pixels, the Euclidean length of a skeleton pixel is different (Supplementary Fig. 9).

Microscopic space to anatomical space. To account for tissue shrinkage (Supplementary Fig. 9), which is inherent to DISCO clearing, we carried out an experiment to measure the degree of shrinkage. Before clearing, we imaged the brains of three live BALB/c mice by MRI and calculated each brain's average volume, through precise manual segmentation by an expert. Next, we cleared three BALB/c brains, processed them with VesSAP and measured the total brain volume with atlas alignment. We report an average volume of $423.84 \pm 2.04\,mm^3$ for the live mice and $255.62 \pm 6.57\,mm^3$ for the cleared tissue. This corresponds to a total volume shrinkage of 39.69%. We applied this as a correction factor for the volumetric information (for example, for brain regions).

Similarly to previous studies, shrinkage was uniform in all three dimensions. This is important when considering shrinkage in one dimension, as needed to account for the shrinkage in centerlines and radii. The one-dimensional correction factor K_L then corresponds to the cube root of the volumetric correction factor K_V.

Accounting for these factors, we calculated the vessel length per volume (Z) in cleared ($Z_{cleared}$) and real (Z_{real}) tissue in equation (5), where $N_{V,vox}$ is the number of total voxels in the reference volume and $N_{Cl,vox}$ is the number of centerline voxels in the image volume:

$$Z_{cleared} = \frac{N_{Cl,vox}}{N_{V,vox}} \times \varepsilon_{Cl} \quad Z_{real} = \frac{N_{Cl,vox}}{N_{V,vox}} \times \varepsilon_{Cl} \times \frac{\kappa_L}{\kappa_V} \quad (5)$$

Similarly, we calculated the bifurcation density (B) in cleared and real tissue in equation (6), where $N_{Bif,vox}$ is the number of bifurcations in the reference volume:

$$B_{cleared} = \frac{N_{Bif,vox}}{N_{V,vox}} \quad B_{real} = \frac{N_{Bif,vox}}{N_{V,vox}} \times \frac{1}{\kappa_V} \quad (6)$$

Please note that the voxel spacing of $3\,\mu m$ has to be taken into consideration when reporting features in SI units.

Inter-annotator experiment for features. To estimate the error in VesSAP's feature quantification, we extracted the features on a labeled test set of five volumes of $500 \times 500 \times 50$ pixels. When comparing to the gold-standard label, we calculated errors (disagreements) of 8.21% for the centerlines, 13.18% for the number of bifurcation points and 16.33% for the average radius. To compare VesSAP's extracted features to human-level annotation, we implemented an inter-annotator experiment. For this experiment, we had four annotators label the vessels and radii in two volumes of $500 \times 500 \times 50$ pixels by using ImageJ and ITK-snap. Finally, we calculated the agreement of the extracted features between all annotators and compared to the gold-standard labeling.

We calculated this for each of the volumes and found an average error (disagreement) of 34.62% for the radius, 25.20% for the bifurcation count and 12.55% for the centerline length.

The agreement between the VesSAP output and the gold standard was higher than the average agreement between the annotators and the gold standard. This difference underlines the quality and reproducibility of VesSAP's feature extraction.

Registration to the reference atlas. We used the average template, the annotation file and the latest ontology file (Ontology ID: 1) of the current Allen mouse brain atlas: CCFv3 201710. Then, we scaled the template and the annotation file up from 10 to $3\,\mu m^3$ to match our reconstructed brain scans and multiplied the left side of the (still symmetrical) annotation file by −1 so that the labels could be later assigned to the corresponding hemispheres. Next, the average template and 3D vascular datasets were downsampled to 10% of their original size in each dimension to achieve reasonably fast alignment with the elastix toolbox[50] (v4.9.0). For the sake of the integrity of the extracted features, we aligned the template to each of the brain scans individually by using a two-step rigid and deformable registration (B-spline; optimizer, AdaptiveStochasticGradientDescent; metric, AdvancedMattesMutualInformation; grid spacing in physical units, 90; in the VesSAP repository, we host the log and parameter files for each brain scan) and applied the transformation parameters to the full-resolution annotation volume

(3-μm resolution). Subsequently, we created masks for the anatomical clusters on the basis of the current Allen brain atlas ontology.

Statistical analysis of features. Data collection and analysis were not performed with blinding to the strains. Data distribution was assumed to be normal, although this was not formally tested. All data values of the descriptive statistics are given as mean ± s.e.m. unless stated otherwise. Data were analyzed with standardized effect size indices (Cohen's d)[51] to investigate differences in vessel length, number of bifurcation points and radii between brain areas across the three mouse strains ($n = 3$ mice per strain). Descriptive statistics were evaluated across brain regions in the pooled ($n = 9$) dataset.

Data visualization. All volumetric datasets were rendered with Imaris, Vision4D and ITK-snap.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
VesSAP data are publicly hosted at http://DISCOtechnologies.org/VesSAP and include original scans and registered atlas data.

Code availability
VesSAP codes are publicly hosted at http://DISCOtechnologies.org/VesSAP and include the imaging protocol, trained algorithms, training data and a reference set of features describing the vascular network in all brain regions. Additionally, the source code is hosted on GitHub (https://github.com/vessap/vessap) and on the executable platform Code Ocean (https://doi.org/10.24433/CO.1402016.v1)[52]. Implementation of external libraries is available on request.

References
41. Breckwoldt, M. O. et al. Correlated magnetic resonance imaging and ultramicroscopy (MR-UM) is a tool kit to assess the dynamics of glioma angiogenesis. *eLife* **5**, e11712 (2016).
42. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. Preprint at https://arxiv.org/abs/1412.6980 (2017).
43. Abadi, M. et al. in *12th USENIX Symposium on Operating Systems Design and Implementation* 265–283 (2016).
44. Hoo-Chang, S. et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **35**, 1285–1298 (2016).
45. Schneider, M., Hirsch, S., Weber, B., Székely, G. & Menze, B. H. Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters. *Med. Image Anal.* **19**, 220–249 (2015).
46. Paetzold, J. C. et al. clDice—a novel connectivity-preserving loss function for vessel segmentation. in *Medical Imaging Meets NeurIPS 2019 Workshop* (2019).
47. Lee, T. C., Kashyap, R. L. & Chu, C. N. Building skeleton models via 3-D medial surface axis thinning algorithms. *Graph. Models Image Process.* **56**, 462–478 (1994).
48. Rempfler, M. et al. Reconstructing cerebrovascular networks under local physiological constraints by integer programming. *Med. Image Anal.* **25**, 86–94 (2015).
49. Marchesi, V. T. Alzheimer's dementia begins as a disease of small blood vessels, damaged by oxidative-induced inflammation and dysregulated amyloid metabolism: implications for early detection and therapy. *FASEB J.* **25**, 5–13 (2011).
50. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. P. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* **29**, 196–205 (2009).
51. Cohen, J. The effect size index: d. *Stat. Power Anal. Behav. Sci.* **2**, 284–288 (1988).
52. Paetzold, J. C. & Tetteh, G. VesSAP: machine learning analysis of whole mouse brain vasculature. *Code Ocean* https://doi.org/10.24433/CO.1402016.v1 (2020).

Author contributions
M.I.T. performed the tissue processing, clearing and imaging experiments. M.I.T. and K.T.-V. developed the whole-brain staining protocol. M.I.T. stitched and assembled the whole-brain scans. V.E. and J.C.P. generated the synthetic vascular training dataset. J.C.P., G.T. and O.S. developed the deep learning architecture and trained the models. J.C.P. and S.S. performed the quantitative analyses. M.I.T. annotated the data. M. Düring and M. Dichgans helped with data interpretation. B.M., M.P. and G.T. provided guidance in developing the deep learning architecture and helped with data interpretation. A.E., M.I.T., B.M. and J.C.P. wrote the manuscript. All authors edited the manuscript. A.E. initiated and led all aspects of the project.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41592-020-0792-1.

Correspondence and requests for materials should be addressed to B.M. or A.E.

Peer review information Nina Vogt was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.

nature research

Corresponding author(s):	Ali Ertürk, Bjoern Menze
Last updated by author(s):	Feb 14, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ImspectorPro [v 5.1] was used for collecting light-sheet images, Zen black [v 10.0] for confocal images and Nucline [v 3.0] for MRI data.
Data analysis	Python [v 2.7 & 3.6], MATLAB [v 9.3], elastix [v 4.9], ITK-SNAP [v 3.6], ImageJ [v 1.52p], Imaris [v 9.3] and Vision4D [v 3.1] were used. Custom codes was used in the study were used for vascular segmentation, mapping and quantification. The description of the software is available in the Method section of the manuscript. The custom codes are publicly available on http://DISCOtechnologies.org/VesSAP

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are publicly available on http://DISCOtechnologies.org/VesSAP

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes for this reproducibility study were determined based on the literatures in the fields. Statistics were not used to predetermine sample sizes.
Data exclusions	No animals were excluded from the study. Specific brain regions in Figure 5a-d were excluded based on visual inspection (zero vascular length), because some regions were damaged and distorted during dissection so that vessels in these regions were not annotated.
Replication	The protocols in the study were replicated successfully more than 5 times in independent experiments. Tissue-clearing, imaging, vascular segmentation, mapping and quantification procedures were performed in nine mouse brains.
Randomization	Within each strain, animals were randomly selected.
Blinding	No blinding was done because knowledge of experimental conditions were required during data collection.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	This study used male mice from the C57BL/6J, CD1 and BALB/c strains at the age of 3 months.
Wild animals	The study did not involve wild animals.
Field-collected samples	This study did not involve a field-collected samples.
Ethics oversight	The experiments were done after approval of the ethical review board of the government of Upper Bavaria (Regierung von Oberbayern, Munich, Germany), and in accordance with the European directive 2010/63/EU for animal research.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	No functional MRI (fMRI) data were acquired.
Design specifications	None.
Behavioral performance measures	None.

Acquisition

Imaging type(s)	Structural MRI.
Field strength	3T
Sequence & imaging parameters	An average of six T1 scans (3D Gradient Echo sequence with TE/TR/flip angle=5.12/14.88/10, a slice thickness = 0.2 mm and 0.1667x0.1667 mm in plane resolution) was created to get a high resolution structural image of the brain.
Area of acquisition	Whole-brain in-vivo.

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	None.
Normalization	No normalization was done to ensure the precise anatomical size measurements of the tissue.
Normalization template	None.
Noise and artifact removal	None.
Volume censoring	No volume censoring was performed.

Statistical modeling & inference

Model type and settings	None.
Effect(s) tested	None.

Specify type of analysis: ☒ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	None.
Correction	None.

Models & analysis

n/a | Involved in the study
☒ ☐ Functional and/or effective connectivity
☒ ☐ Graph analysis
☒ ☐ Multivariate modeling or predictive analysis