PLOS ONE

RESEARCH ARTICLE

Deep convolutional neural networks for segmenting 3D *in vivo* multiphoton images of vasculature in Alzheimer disease mouse models

Mohammad Haft-Javaherian[1], Linjing Fang[1], Victorine Muse[1], Chris B. Schaffer[1], Nozomi Nishimura[1], Mert R. Sabuncu[1,2]*

1 Nancy E. and Peter C. Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, United States of America, **2** School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, United States of America

* msabuncu@cornell.edu

Citation: Haft-Javaherian M, Fang L, Muse V, Schaffer CB, Nishimura N, Sabuncu MR (2019) Deep convolutional neural networks for segmenting 3D *in vivo* multiphoton images of vasculature in Alzheimer disease mouse models. PLoS ONE 14(3): e0213539. https://doi.org/10.1371/journal.pone.0213539

Editor: Ulas Bagci, University of Central Florida (UCF), UNITED STATES

Received: October 18, 2018

Accepted: February 22, 2019

Published: March 13, 2019

Data Availability Statement: We have made all data used for the analyses presented in our paper publicly available at: https://hdl.handle.net/1813/59221.

Funding: This work was supported by the European Research Council grant 615102 (NN), the National Institutes of Health grant AG049952 (CS), the National Institutes of Health grants R01LM012719 and R01AG053949 (MS), the National Science Foundation Cornell NeuroNex Hub

Abstract

The health and function of tissue rely on its vasculature network to provide reliable blood perfusion. Volumetric imaging approaches, such as multiphoton microscopy, are able to generate detailed 3D images of blood vessels that could contribute to our understanding of the role of vascular structure in normal physiology and in disease mechanisms. The segmentation of vessels, a core image analysis problem, is a bottleneck that has prevented the systematic comparison of 3D vascular architecture across experimental populations. We explored the use of convolutional neural networks to segment 3D vessels within volumetric *in vivo* images acquired by multiphoton microscopy. We evaluated different network architectures and machine learning techniques in the context of this segmentation problem. We show that our optimized convolutional neural network architecture with a customized loss function, which we call *DeepVess*, yielded a segmentation accuracy that was better than state-of-the-art methods, while also being orders of magnitude faster than the manual annotation. To explore the effects of aging and Alzheimer's disease on capillaries, we applied *DeepVess* to 3D images of cortical blood vessels in young and old mouse models of Alzheimer's disease and wild type littermates. We found little difference in the distribution of capillary diameter or tortuosity between these groups, but did note a decrease in the number of longer capillary segments (>75μm) in aged animals as compared to young, in both wild type and Alzheimer's disease mouse models.

Introduction

The performance of organs and tissues depend critically on the delivery of nutrients and removal of metabolic products by the vasculature. Blood flow deficits due to disease related factors or aging often leads to functional impairment [1]. In particular, the brain has essentially no energy reserve and relies on the vasculature to provide uninterrupted blood perfusion [2].

grant (1707312, MS and CS) and the National Science Foundation (1748377 to MS).

Competing interests: The authors have declared that no competing interests exist.

Multiple image modalities can be used to study vascular structure and dynamics, each offering tradeoffs between the smallest vessels that can be resolved and the volume of tissue that can be imaged. Recent work with several modalities, including photoacoustic microscopy [3], optical coherence tomography [4], and multiphoton microscopy (MPM) [5], enable individual capillaries to be resolved in 3D over volumes approaching 1 mm^3 in living animals. The analysis of such images is one of the most critical and time-consuming tasks of this research, especially when it has to be done manually.

For example, in our own work we investigated the mechanisms leading to reduced brain blood flow in mouse models of Alzheimer's disease (AD), which required extracting topology from capillary networks each with $\sim$1, 000 vessels from dozens of animals. The manual tracing of these networks required $\sim$40$\times$ the time required to acquire the images, greatly slowing research progress [6]. The labor involved in such tasks limits our ability to investigate the vital link between capillary function and many different diseases. Many studies have shown anatomical and physiological differences in microvasculature associated both with age and AD, such as changes in composition of large vessel walls' smooth muscles [7], increased collagen VI in microvascular basement membranes and their thickening in AD [8], and age-associated reduction of microvascular plasticity and the ability of the vessels to respond appropriately to changes in metabolic demand [9].

In this paper, we consider the segmentation of vessels, a core image analysis problem that has received considerable attention [10, 11]. As in other segmentation and computer vision problems, in recent years deep neural networks (DNNs) have offered state-of-the-art performance [12]. DNN approaches often rely on formulating the problem as supervised classification (or regression), where a neural network model is trained on some (manually) labeled data. For a survey on deep learning in medical image analysis, see a recent review by Litjens et al. [12].

Here, we explore the use of a convolutional neural network (CNN) to segment 3D vessels within volumetric *in vivo* MPM images. *In vivo* MPM imaging of blood vessels has the advantage that it captures the size and shape of vessels without introducing artifacts from postmortem tissue processing. However, blood flow generates features which must be accommodated in the vessel segmentation. We conduct a thorough study of different network architectures and machine learning techniques in the context of this segmentation problem. We apply the final model, which we call *DeepVess*, on image stacks of cortical blood vessels in mouse models of AD and wild type (WT) littermates. Our experimental results show that *DeepVess* yields segmentation accuracy that is better than current state-of-the-art, while being orders of magnitude faster than the manual annotation (20-30 hours manual work vs. 10 minutes computation time). The segmentation method developed in this work provides robust and efficient analysis which enabled us to quantify and compare capillary diameters and other vascular parameters from *in vivo* cortex images across multiple animals, with varying age as well as across WT mice and AD models.

Related work

Blood vessel segmentation is one of the most common and time-consuming tasks in biomedical image analysis. This problem can either be approached in 2D or 3D, depending on the specifics of the application and analytic technique. The most established blood vessel segmentation methods are developed for 2D retinography [13] and 3D CT/MRI [11].

Among segmentation methods, region-based methods are well-known for their simplicity and low computational cost [14]. For example, Yi et al. [15] developed a 3D region growing vessel segmentation method based on local cube tracking. In related work, Mille et al. [16]

used a 3D parametric deformable model based on the explicit representation of a vessel tree to generate centerlines. In recent years, these traditional segmentation methods have become less popular and are considered to be limited in comparison to deep learning methods, because they require handcrafted filters, features, or logical rules and often yield lower accuracy.

Today, in problems that are closely related to ours, various deep learning techniques dominate state-of-the-art. For instance, in a recent Kaggle challenge for diabetic retinopathy detection within color fundus images, deep learning was used by most of the 661 participant teams, including the top four teams. Interestingly, those top four methods surpassed the average human accuracy. Subsequently, Gulshan et al. [17] adopted the Google Inception V3 network [18] for this task and reached the accuracy of seven ophthalmologists combined. For retinal blood vessel segmentation, Wu et al. [19] used a CNN-based approach to extract the entire connected vessel tree. Fu et al. [20] proposed to add a conditional random fields (CRF) to post-process the CNN segmentation output. They further improved their method by replacing the CRF with a recurrent neural network (RNN), which allows them to train the complete network in an end-to-end fashion [21]. Further, Maninis et al. [22] addressed retinal vessel and optic disc segmentation problems using one CNN network and could surpass the human expert.

There are 3D capillary image datasets in mice [14] and human [23] that were segmented using traditional segmentation methods and have illustrated the scientific value of such information, but few such datasets are available.

To the best of our knowledge, there are only two studies that used deep learning for our problem: vascular image analysis of multi-photon microscopy (MPM) images. The first one is by Teikari et al. [24] who proposed a hybrid 2D-3D CNN architecture to produce state-of-the-art vessel segmentation results in 3D microscopy images. The main limitation of their method was the use of 2D convolutions and 2D conditional random fields (CRF)s, which restrict the full exploitation of the information along the third dimension. The second study was conducted by Bates et al. [25], where the authors applied a convolutional long short-term memory RNN to extract 3D vascular centerlines of endothelial cells. Their approach was based on the U-net architecture [26], which is a well-known fully convolutional network [27] widely used for biomedical image segmentation. Bates and colleagues achieved state-of-the-art results in terms of centerline extraction; nevertheless, they reported that certain vessels in the images were combined in the automatic segmentation. Finally, we consider the 3D U-Net [28], which is the volumetric version of the U-net architecture [26] and is regarded by many as state-of-the-art for microscopy image segmentation problems.

Data and methods

The proposed vasculature segmentation method for 3D *in vivo* MPM images, *DeepVess*, consists of (i) pre-processing to remove *in vivo* physiological motion artifacts due to respiration and heartbeat, (ii) applying a 3D CNN for binary segmentation of the vessel tree, and (iii) post-processing to remove artifacts such as network discontinuities and holes.

Data

Animals. All animal procedures were approved by the Cornell University Institutional Animal Care and Use Committee and were performed under the guidance of the Cornell Center for Animal Resources and Education. We used double transgenic mice (B6.Cg-Tg (APPswe, PSEN1dE9) 85Dbo/J, referred to as APP/PS1 mice) that express two human proteins associated with early onset AD, a chimeric mouse/ human amyloid precursor protein (Mo/HuAPP695swe) and a mutant human presenilin1 (PS1-dE9), which is a standard model of AD

and typically develops amyloid-beta plaque deposition around 6 months of age [29]. Littermate WT mice (C57BL/6) served as controls. Animals were of both sexes and ranged in age from 18 to 31 weeks for young mice and from 50 to 64 weeks for the old mice (6 WT and 6 AD at each age, for a total of 24 mice).

***In vivo* imaging of cortical vasculature.** We use a locally-designed multiphoton microscope [30] for *in vivo* imaging of the brain vasculature. Glass-covered craniotomies were prepared over parietal cortex, as described previously [6, 31, 32]. For cranial window implantation and imaging, mice were anesthetized with 3% isoflurane and then maintained on 1.5% isoflurane in 100% oxygen. Mice were injected with 0.05 mg/100g of mouse weight glycopyrrolate (Baxter Inc.) or 0.005 mg/100g atropine (intramuscular 54925-063-10, Med-Pharmex Inc.). At time of surgery as well as 1 and 2 days after mice received 0.025 mg/100g dexamethasone (subcutaneous 07-808-8194, Phoenix Pharm Inc.), and 0.5 mg/100g ketoprofen (intramuscular, Zoetis Inc.). Bupivacaine (0.1 ml, 0.125%, Hospira Inc.) was subcutaneously injected at the incision site. Animals were injected with 1 ml/100g mouse 5% (w/v) glucose in normal saline subcutaneously every hour during imaging and surgery. Body temperature was maintained at 37 C with a feedback-controlled heating blanket (40-90-8D DC, FHC). Mice were euthanized with pentobarbital overdose after their last imaging session.

We waited at least three weeks after the surgery before imaging to give time for the mild surgically-induced inflammation to subside. Windows typically remained clear for as long as 20 weeks. This technique allows us to map the architecture of the vasculature throughout the top 500 μm of the cortex. Briefly, the blood plasma of an anesthetized mouse was labeled with an intravenous injection of Texas Red labeled dextran (70 KDa, Life Technologies). The two-photon excited fluorescence intensity was recorded while the position of the focus of a femtosecond laser pulse train was scanned throughout the brain, providing a three-dimensional image of the vasculature [30]. Imaging was done using 800-nm or 830-nm, 75-fs pulses from a Ti:Sapphire laser oscillator (MIRA HP, pumped by a Verdi-V18, or Vision S, Coherent). Lasers were scanned by galvonometric scanners and focused into the sample using a 1.0 NA, 20X water-immersion objective lens (Carl Zeiss, Inc.). Image stacks were acquired with 645/45 nm (center wavelength/bandwidth) bandpass filters. The ScanImage software package [33] was used to control the whole system. Image stacks were taken with a range of magnifications resulting in lateral voxel sizes from 0.45 to 1.71 μm/pixel, but always 1 μm in the axial direction.

Expert annotation. We implemented a protocol to facilitate the manual 3D segmentation task using ImageJ, an open-source image processing software package [34] (supplementary material). Two people, one expert and one less experienced, each manually segmented a motion artifact corrected (see below), $256 \times 256 \times 200$ voxels ($292 \times 292 \times 200$ μm^3) image from an AD mouse, independently, which took about 20 and 30 hours, respectively. The second annotator was trained by the expert and then had several months of practice prior to performing this task. These data were used to estimate inter-human segmentation variation. We treated the expert labels as the "gold standard" segmentation and used the second annotator's labels to compare variability in manual segmentation. All other comparisons were made with respect to the gold standard segmentation as the ground truth. This dataset was divided into independent (i.e., non-overlapping) training, validation, and testing sub-parts (50%-25%-25%), all spanning the entire depth of the stack. The training and validation datasets were used in the optimization of CNN architectures, while the test dataset was kept unused until the end of our architecture design optimization process and used for the final unbiased evaluation. We repeated this process 4 times, by varying the test data and thus effectively conducting 4-fold cross-validation. We note that architecture optimization was only done in the first fold. Additionally, six independent 3D images (different mice and different voxel size) acquired by Cruz

Herneández et al. [6] were labeled by an expert to examine the generalization of *DeepVess*. The detailed properties of these images are in S3 Table. With this paper, we also have made all images and expert annotations publicly available at: https://doi.org/10.7298/X4FJ2F1D.

Preprocessing

Motion artifacts caused by physiological movements are one of the major challenges for 3D segmentation of *in vivo* MPM images. Furthermore, global linear transformation models cannot compensate for the local nonuniform motion artifacts, for example, due to a breath occurring part way though the raster scanning for an MPM image. In this study, we adopted the non-rigid non-parametric diffeomorphic demons image registration tool implemented based on the work of Thirion [35] and Vercauteren et al. [36]. Our approach is to register each slice to the previous slice, starting from the first slice as the fixed reference. The diffeomorphic demons algorithm aims to match the intensity values between the reference image and deformed image, where cost is computed as the mean squared error. The smoothness prior on the deformation field is implemented via an efficient Gaussian smoothing of gradient fields, and invertibility is ensured via concatenation of small deformations. This kernel is effectively encouraging the deformation field to be smooth, thus regularizing the ill-posed non-linear registration problem. Based on our experiments, a Gaussian kernel with the standard deviation of 1.3 was chosen for the regularization of the registration algorithm. Next, in our pre-processing steps, the 1-99% range of the image intensities in the input image patch were linearly mapped between 0 and 1, and the extreme 1% of voxels were clipped at 0 and 1. This step, we found, helps with generalizing the model to work well with images taken from other MPM platforms by adapting normalization parameters to the acquisition systems and image statistics utilizing most of the intensity rang. To facilitate comparison between different datasets, image volumes were resampled to have 1 μm^3 voxel for comparisons.

Convolutional neural network architectures

Our aim in this work is to design a system that takes an input stack of images (in 3D) and produces a segmentation of vessels as a binary volume of the same size. For this task, as we elaborate below, we explored different CNN architectures using validation performance as our guiding metric. Our baseline CNN architecture starts with a 3D input image patch (tile), which has 33 × 33 × 5 voxels (in x, y, and z directions). The first convolution layer uses a 7 × 7 × 5 voxel kernel with 32 features to capture 3D structural information within the neighborhood of the targeted voxel. The output of this layer, 32 nodes of 27 × 27 × 1 voxel images, enter a max pooling layer with a 2 × 2 kernel and 2 × 2 strides. Another convolution layer with 5 × 5 × 1 kernel and 64 features, followed by a similar max pooling layer are then applied before the application of the fully connected dense layer with 1024 hidden nodes and dropout [37] with a probability value of 50%. The output is a two-node layer, which represents the probability that the pixel at the center of the input patch belongs to tissue vs. vessel. The CNN takes an input 3D patch and produces a segmentation label for the central voxel. All the convolution layers have a bias term and rectified linear unit (ReLU) as the element-wise nonlinear activation function. Starting from this baseline CNN architecture, we optimized the network architecture hyperparameters with a greedy algorithm.

Different kernel sizes for the 3D convolution layers were explored in our experiments. Note that each choice in the architecture parameters (including the kernel size) corresponded to a different input patch size. As the validation results summarized in S1 Table indicate, the best performing baseline architecture had an input patch size of 33 × 33 × 7. Based on this result we chose an input patch size of 33 × 33 × 7 as the optimal field of view (FOV) for segmentation.

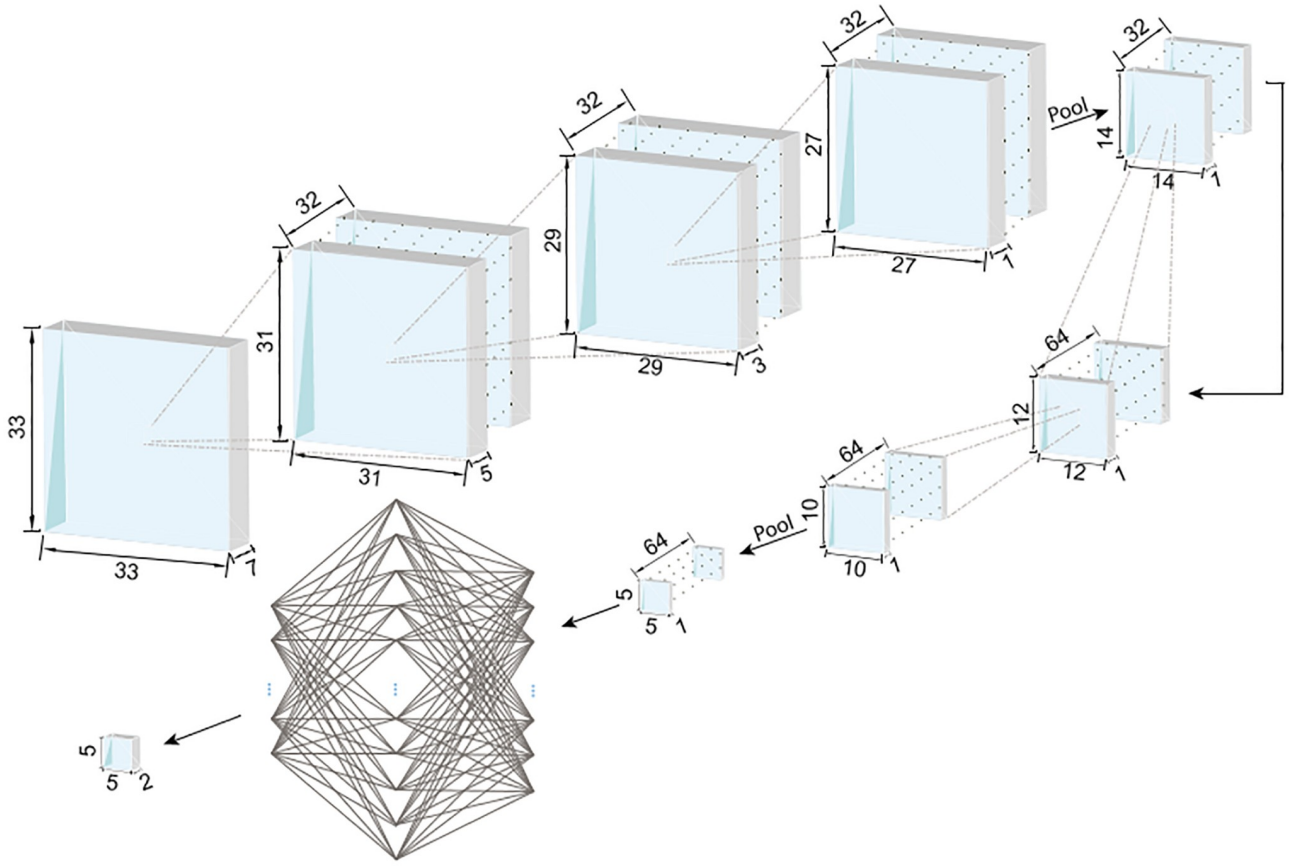

Fig 1. The optimal 3D CNN architecture. The field of view (FOV), i.e. the input patch size, is 33 × 33 × 7 voxels and the output is the segmentation of the 5 × 5 × 1 patch (region of interest, ROI) at the center of the patch. The convolution kernels are 3 × 3 × 3 voxels for all the layers and ReLU is used as the element-wise nonlinear activation function. The first three convolution layers have 32 channels and are followed by pooling. The second three convolution layers have 64 channels. The output of convolution layers is 5 × 5 × 1 voxels with 64 channels, which is fed to a fully connected neural network with a 1024-node hidden layer. The final result has 5 × 5 × 1 voxels with two channels representing the probability of the foreground and background label associations.

https://doi.org/10.1371/journal.pone.0213539.g001

We then explored the effect of the number of convolutional and max pooling layers. As summarized in S2 Table, the best architecture had three 3D convolution layers with a 3 × 3 × 3 voxel kernel, a max pooling layer, followed by two convolution layers with a 3 × 3 voxel kernel, and a max pooling layer. The output of the last max pooling layer is reshaped to a fully-connected layer followed by a 1024-node fully-connected layer and the last fully-connected layer, which is reshaped to the output patch size. Note that there is no difference in spatial resolution (i.e., voxel dimensions) between the input and output patches.

Finally, we investigated the performance for different output patch sizes, ranging from 1 voxel to 5 × 5 × 5 voxels and found that performance was improved further when the output is the segmentation of the central 5 × 5 × 1 patch and not just a single voxel. A larger output area has the advantage of accounting for the structural relationship between adjacent voxels in their segmentation. The optimal CNN architecture scheme is shown in Fig 1.

Performance metrics

There are different performance metrics to compare agreement between an automated segmentation method and a "ground truth" (GT) human annotation. In the context of binary

segmentation, the foreground (F) will be the positive class, and the negative class will correspond to the background (B). Therefore, true positive (TP) can be measured as the total number of voxels where both the automatic and human segmentation labels are foreground. True Negative (TN), False Positive (FP) and False Negative (FN) can be defined in a similar fashion.

Based on these, we can compute *sensitivity* and *specificity*. For example, sensitivity is the percentage of GT foreground voxels that are labeled by the automatic segmentation (ASeg) correctly. Mathematically, we have:

$$sensitivity = P(y = F|GT = F) = \frac{TP}{TP + FN} \tag{1}$$

$$specificity = P(y = B|GT = B) = \frac{TN}{TN + FP} \tag{2}$$

The *Dice coefficient* (DC), *Jaccard index* (JI), and *modified Hausdorff distance* (MHD) are another set of commonly used segmentation performance metrics. JI is defined as the ratio between the number of voxels labeled as foreground by both GT and ASeg, to the total number of voxels that are called foreground by either GT and ASeg. DC is very similar to JI, except it values TP twice as much as FP and FN. JI and DC are useful metrics when the number of the foreground voxels is much less than background and the detection accuracy of the foreground voxels is more important compared to background voxel detection, which is the case for 3D imaging of vasculature.

$$JI = P(y = F \cap GT = F \mid y = F \cup GT = F) = \frac{TP}{TP + FP + FN} \tag{3}$$

$$DC = \frac{2 \times JI}{1 + JI} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{4}$$

On the other hand, MHD [38] quantifies accuracy in terms of distances between boundaries, which might be appropriate when considering tubular structures. For each boundary point in image A($a \in \mathcal{A}$), the closest Euclidean distance ($d(a, b) = ||a - b||_2$) to any boundary point inside image B($b \in \mathcal{B}$) is first calculated, $d(a, \mathcal{B}) = min_{b\in\mathcal{B}} \ ||a - b||_2$). This is then averaged over all boundary points in $\mathcal{A}$: $\frac{1}{N_a}\sum_{a\in\mathcal{A}} d(a, \mathcal{B})$ [39]. MHD is then defined as:

$$MHD = max\left[\frac{1}{N_a}\sum_{a\in\mathcal{A}} d(a, \mathcal{B}), \ \ \frac{1}{N_b}\sum_{b\in\mathcal{B}} d(b, \mathcal{A})\right] \tag{5}$$

$$d(a, \mathcal{B}) = min_{b\in\mathcal{B}} \ ||a - b||_2 \tag{6}$$

Note that in the segmentation setting, **A** and **B** can represent the foreground boundaries in the automatic and GT segmentations, respectively. Finally, we can compute the MHD on centerlines instead of boundaries, a metric we call MHD-CL.

Training and implementation details

In training our segmentation algorithms, we used a customized cross-entropy loss function designed for our highly unbalanced datasets (where foreground voxels comprise only a small fraction of the volume), measured over all voxels but TN ($i \in \{TP, FP, FN\}$), defined as:

$$Loss = \sum_{i\in\{TP,FP,FN\}} -[y_i \ \log(p_i) + (1 - y_i)\log(1 - p_i)] \tag{7}$$

y_i is the GT label and p_i is the model's output as the probability of the target voxel i belonging to the foreground. Note that in Eq (7), true negative voxels have no contribution, effectively reducing the influence of the dominant background. We trained our model using *Adam* stochastic optimization [40] with a learning rate of 10^{-4} for 100 epochs during architecture exploration and a learning rate of 10^{-6} for 30,000 epochs during the fine tuning of model parameters for the proposed architecture with mini-batch size of 1000 samples (based on GPU memory constraints and results of our experiments with smaller mini-batch size, which did not improve the optimization results). The fine tuning took one month on one NVIDIA TITAN X GPU. We implemented our models in Python using Tensorflow™ [41].

Post-processing

CNN segmentation results contain some segmentation artifact such as holes inside the vessels, rough boundaries, or isolated small objects. In order to remove these artifacts, the holes within the vessels were filled. This was followed by application of a 3D mean filter with a $3 \times 3 \times 3$ voxel kernel and the removal of small foreground objects, e.g. smaller than 100 voxels. This result was used to compare to the gold standard.

Analysis of vasculature centerlines

To characterize the cortical vasculature of the experimental animals, we identified capillary segments by calculating centerlines from the segmented image data. Our centerline extraction method includes dilation and thinning operations, in addition to some centerline artifact removal steps. The binary segmentation image was first thinned using the algorithm developed by Lee et al. [42]. The result was then dilated using a spherical kernel with a radius of 5-voxels to improve the vessel connectivity, which was followed by mean filtering with a $3 \times 3 \times 3$ voxel kernel and removing holes from each cross section. Next, a thinning step was applied again to obtain the new centerline result. The original segmented image was dilated using a spherical kernel with a radius of 1-voxel to act as the mask for the centerlines with the goal of improving the centerline connectivity. The following rules were applied to the resulting centerlines repeatedly until no further changes could be done. A vessel is a segment between two bifurcations.

1. Remove any vessels with one end not connected to the network (i.e., dead end) and with length smaller than 11 voxels.
2. Remove single voxels connected to a junction.
3. Remove single voxels with no connections.
4. Remove vessel loops with length of one or two voxels.

Finally, the centerline network representation (i.e. nodes, edges, and their properties) was extracted. (The centerline extraction was applied on both manual and automated segmentations.)

Results

We conducted a systematic evaluation of several network architecture parameters in order to optimize segmentation accuracy of images of mouse cortex vasculature from MPM. Features of *in vivo* MPM images include motion artifacts due to respiration and heart beat. Because vessels are visualized by an injection of dye that labels the blood plasma, unlabeled red blood cells appear as dark spots and streaks moving through the vessel lumen (arrows in Fig 2). Images

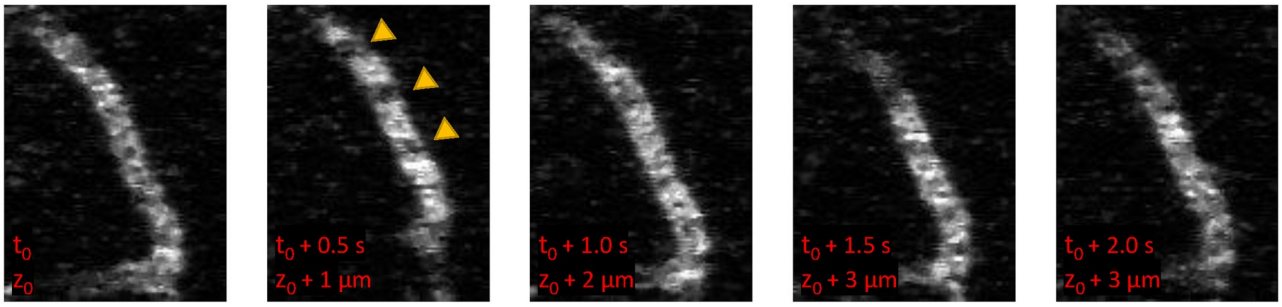

Fig 2. *In vivo* MPM images of a capillary. Because MPM images are acquire by raster scanning, images at different depths (z) are acquired with a time lag (t). Unlabeled red blood cells moving through the lumen cause dark spots and streaks and result in variable patterns within a single vessel.

https://doi.org/10.1371/journal.pone.0213539.g002

are acquired by raster scanning through the tissue and each vessel is captures in several images. The imaging speed has a significant influence on these features and in *in vivo* experiments, imaging is often relatively slow, such that these features become prominent [43]. We emphasize that our exploration was based on performance on the validation dataset and the final results presented reflect the model accuracy on an independent test dataset. The detailed performance results for some of the tested architectures are reported in S1 and S2 Tables. The optimal architecture, *DeepVess*, was trained on the training data until the model accuracy stopped improving and no overfitting was observed (30,000 epochs). S1 Fig shows the JI learning curve over 30,000 epochs, for training, validation, and test datasets. The constant gap between JI of the training and validation datasets, which represent generalization error, confirms that we are not strongly overfitting.

Furthermore, we implemented two state-of-the-art methods [24, 28], and an improved version of the method of Teikari et al. [24], where we changed the 2D convolutional kernels into 3D kernels and inserted a fully connected neural network layer at the end, based on the suggestion in the discussion of their paper. Table 1 summarizes the comparison between the performance of our optimal architecture based on the 4-fold cross-validation results, with and without the post-processing step, comparing to two state-of-the-art methods and a second human annotator to provide a measure of the inter-human variability. These results, as well as S1 Fig demonstrate that *DeepVess* outperforms the state-of-the-art methods [24, 28] in terms of sensitivity, Dice index, Jaccard index, and boundary modified Hausdorff distance; and approaches human performance in terms of Dice and Jaccard. The proposed method does not outperform the benchmarks in specificity, indicating a slightly higher rate of false positive voxels. Yet we note that the relatively lower specificity is still very high (97%).

Table 1. The comparison of our proposed CNN architecture (*DeepVess*), manual annotation by a trained person, and two state-of-the-art methods [24, 28] to the gold standard of the expert human annotation based on the 4-fold cross-validation results. *DeepVess* surpassed both human annotator and two state-of-the-art methods in terms of sensitivity as well as Dice index, Jaccard index, and boundary modified Hausdorff distance, which are the three metrics that are widely used in segmentation.

	Sensitivity	*Specificity*	*Dice*	*Jaccard*	*MHD*
Second human annotator	81.07%	98.70%	82.35%	70.40%	1.50
Original Teikari et al. [24]	62.44%	**98.65%**	69.69%	55.06%	3.20
Çiçek et al. [28]	70.01%	98.21%	72.69%	59.41%	3.55
Improved [24] in this study	69.55%	98.39%	74.03%	59.96%	3.16
DeepVess	89.91%	97.00%	81.62%	69.13%	2.26
DeepVess with post-processing	**89.95%**	97.00%	**81.63%**	**69.15%**	**2.25**

https://doi.org/10.1371/journal.pone.0213539.t001

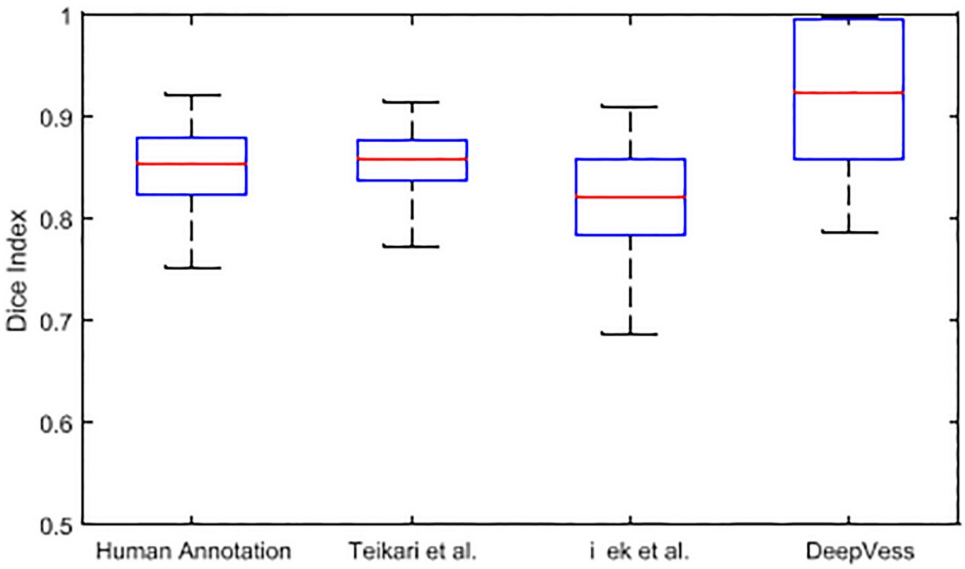

Fig 3. Slice-wise Dice index of *DeepVess* vs. manual annotation by a trained person and the state-of-the-art methods [24, 28] compared to the gold standard of the expert human annotation. The central red mark is the median, and the top and bottom of the box is the third and first quartiles, respectively. The whiskers indicates the range of data. *DeepVess* has higher median value in comparison to the Teikari et al. [24], Çiçek et al. [28], and the human annotator (Wilcoxon signed-rank test, $p = 2.98e-23$, $p = 2.59e-32$, and $p = 2.8e-28$, respectively).

https://doi.org/10.1371/journal.pone.0213539.g003

In MPM, the variation in the signal to noise as a function of imaging depth leads to changes in image quality between image slices. The performance of a segmentation method should therefore be assessed by analyzing slices separately. Fig 3 illustrates the boxplot of slice-wise Dice index values from the x-y planes within the 3D MPM image dataset. *DeepVess* had a higher Dice index values in comparison to the Teikari et al. and the trained annotator's results. However, there was more variation compared to the other two results, which implies the possibility and need for further improvements.

The generalization of the model was studied by testing an independent dataset annotated by our expert consisting of 6 separate 3D MPM images acquired from 1 AD and 5 WT mice (S3 Table) and the results are summarizes in S4 Table. *DeepVess* outperforms both the state-of-the-art methods [24, 28] on the second dataset in terms of sensitivity, Dice index, Jaccard index, and boundary MHD. Similar to the test dataset results, specificity was slightly lower. These results illustrate the generalization of our model on new MPM images with different image quality and captured from different mouse models and with different voxels sizes. Fig 4A illustrates the image intensity and three models overlaid on the image for a cross-section extracted from a 3D image from the independent dataset (S4 Table #1). Fig 4B–4E are magnified version of three cases within Fig 4A. The main sources of failure in the vessel segmentations of 3D *in vivo* MPM images are low SNR at deeper cross-sections (Fig 4C) and unlabeled, moving red blood cells in the vessel lumen, which cause dark spots and streaks (Fig 4B and 4D). The patchy segmentations due to unlabeled red blood cells result in unconnected and isolated vasculature centerlines and network. The *DeepVess* architecture has fully connected layers and thus might be exploiting some spatially varying properties of the signal (as in the variation of contrast as a function of depth) that a fully convolutional architecture such as U-Net might not be able to exploit. Elsewhere, in the absence of such difficulties, all three models segment the vessels largely accurately.

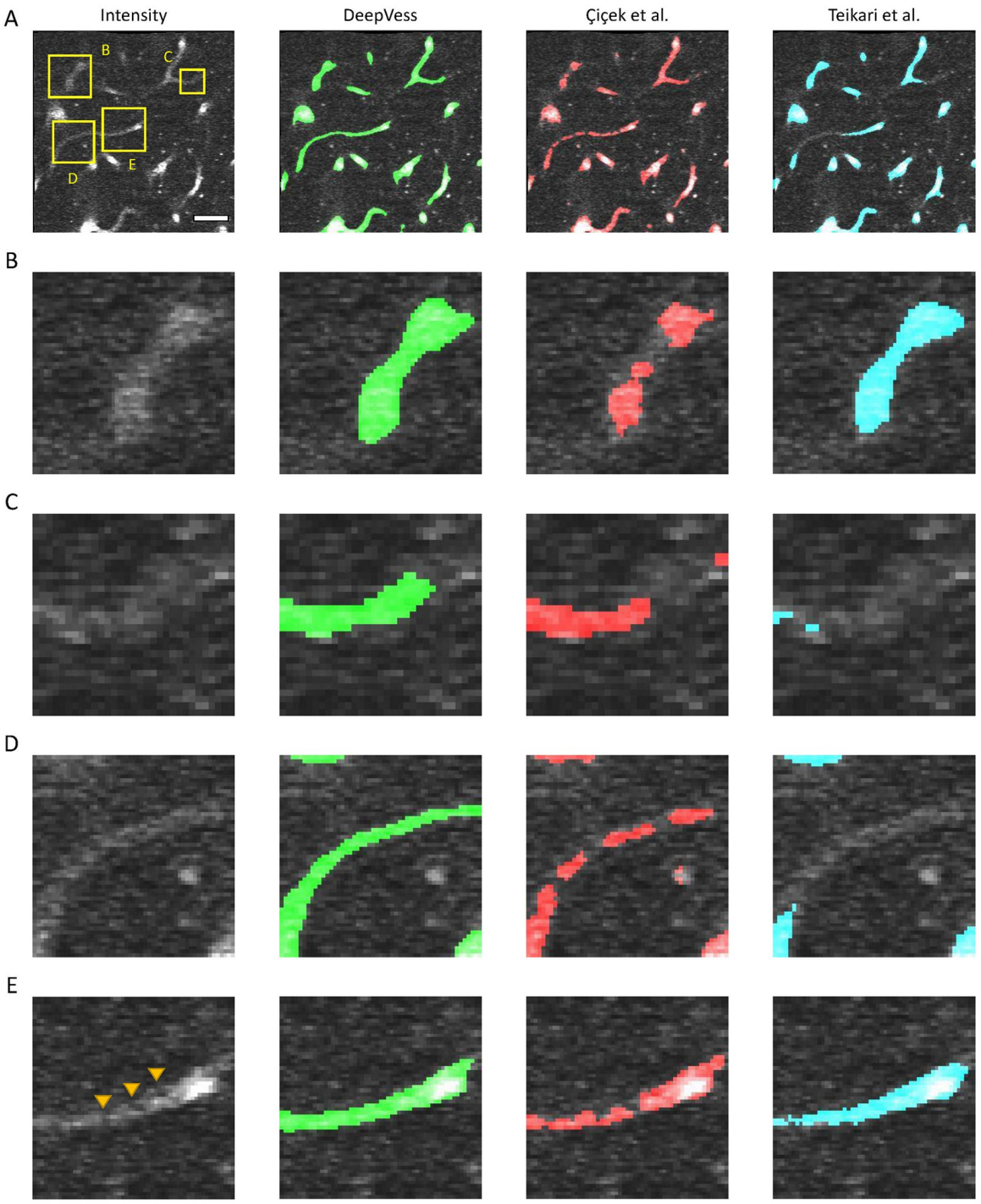

Fig 4. Comparison of *DeepVess* and the state-of-the-art methods [24, 28] in a 3D image cross-section obtained from an independent dataset (S4 Table #1) not used during the training. (A) An image frame with intensity in gray and overlay of segmentation from each method. (B-E) magnified view of four cases within A. The three models overlaid on the complete 3D image is made available online in Supplemental Materials. Scale bar is 50μm.

https://doi.org/10.1371/journal.pone.0213539.g004

We next examined the quality of the vessel centerlines derived from the different segmentations. Using the centerline modified Hausdorff distance (CL MHD) as a centerline extraction accuracy metric, *DeepVess* (CL MHD [*DeepVess*] = 3.03) is substantially better than the state-of-the-art methods (CL MHD [Teikari et al.] = 3.72, CL MHD [Çiçek et al.] = 6.13). But there is still room for improvement in terms of automatic centerline extraction as neither automatic methods yielded scores as good as the trained human annotator (CL MHD [human annotator] = 2.73). In order to test the accuracy of geometrical measurements, the vessel diameter, a sensitive metric, was selected. We measured the diameter of 100 vessels manually by averaging ten 2D measurements per vessel to compare with the *DeepVess's* results (S2 Fig). We observed that there is no significant difference between manually measured diameters and *DeepVess's* results (paired t-test, $n = 100$, $p = 0.34$).

Discussion

The segmentation of 3D vasculature images is a laborious task that slows down the progress of biomedical research and constrains the use of imaging in clinical practice. There has been significant research into tackling this problem via image analysis methods that reduce or eliminate human involvement. In this work, we presented a CNN approach, which surpasses the state-of-the-art vessel segmentation methods [24, 28] as well as a trained human annotator. The proposed algorithm, *DeepVess*, segments 3D *in vivo* vascular MPM images with more than ten million voxels in ten minutes on a single NVIDIA TITAN X GPU, a task that takes 30 hours for a trained human annotator to complete manually.

In order to characterize the performance of *DeepVess*, we compared the automated segmentation to an expert manual segmentation (Fig 5). Here, we visualized three slices with different qualities of segmentation results. The 3D rendering of the mouse brain vasculature shown in Fig 5 indicates the location of these top, middle, and bottom slices representing typical high,

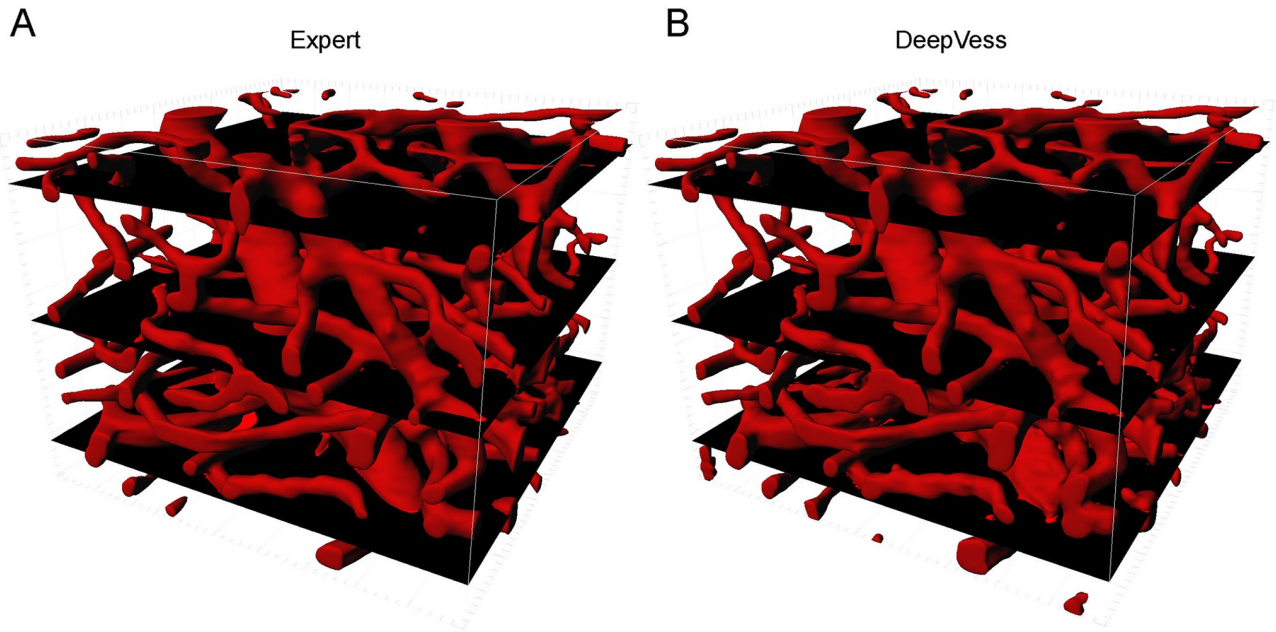

Fig 5. 3D rendering of (A) the expert's manual and (B) *DeepVess* segmentation results. The top, middle, and bottom black plains correspond to the high, medium, and low quality examples, respectively, which are analyzed further in the Discussion (Fig 6). Each volume is 256 × 256 × 200 voxels (292 × 292 × 200 μm^3).

https://doi.org/10.1371/journal.pone.0213539.g005

medium, and low segmentation quality, respectively. Top layers are very similar, and differences are visible at the bottom layers, which have low SNR.

We used 50% dropout during test-time [44] and computed Shannon's entropy for the segmentation prediction at each voxel to quantify the uncertainty in the automated segmentation. Higher entropy represents higher segmentation uncertainty at a particular voxel. The entropy results together with the comparison between *DeepVess* and the expert segmentations for those three planes are illustrated in Fig 6. The left column contains the intensity gray-scale

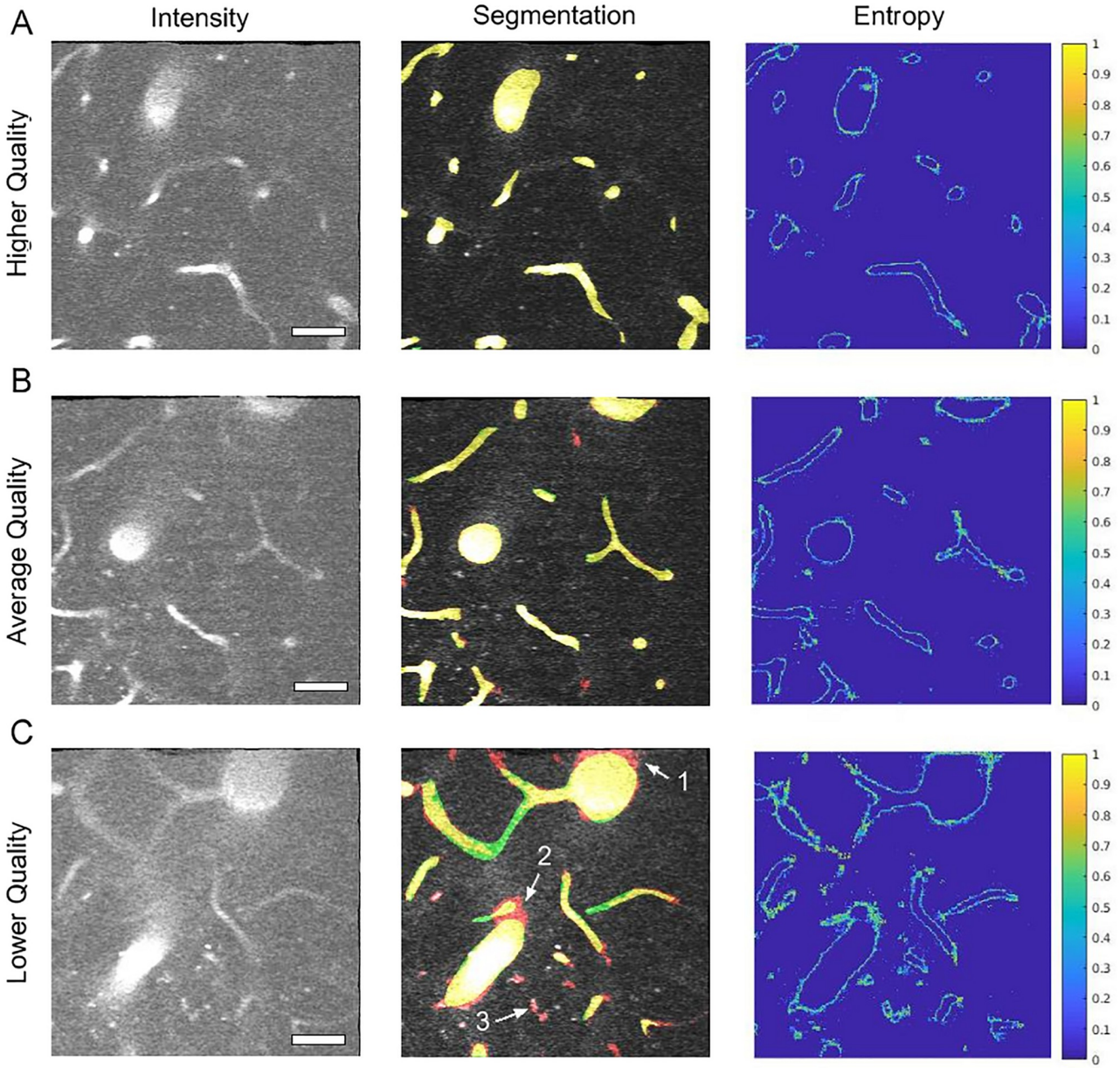

Fig 6. Comparison of *DeepVess* and the gold standard human expert segmentation results in image planes as shown in Fig 5. Imaging is generally higher quality at planes closer to the sample surface. (Left column) Image intensity shown with gray scale after motion artifact removal. The dark spots within the vessels are red blood cells that do not take up the injected dye. (Middle column) Comparison between *DeepVess* (red) and the expert (green) segmentation results overlaid on images. Yellow shows agreement between the two segmentations. (Right column) Shannon entropy, which is a metric of *DeepVess* segmentation uncertainty computed with 50% dropout at test-time [44]. The boundaries of vessels with high entropy values, shown in warmer colors, demonstrate the uncertainty of *DeepVess* results at those locations. Scale bar is 50μm.

https://doi.org/10.1371/journal.pone.0213539.g006

images of these examples. The segmentation results of the *DeepVess* and the expert are superimposed on the original gray-scale image with red (*DeepVess*) and green (the expert), as shown in the middle column. Yellow represents agreement between *DeepVess* and the expert. The right column shows the entropy of each example estimated via test time dropout. We observe that, in general, *DeepVess* has higher uncertainty at the boundaries of vessels. The disagreement with ground truth is also mostly concentrated at the boundaries. Images from deeper within the brain tissue that often have lower image contrast and higher noise levels due to the nature of MPM, suffer from more segmentation errors. These images can often be challenging even for expert humans. Arrows in Fig 6C highlight examples of these difficulties. The error example 1 illustrates the case where the expert ignored bright pixels around the vessel lumen based on their knowledge of the underlying physiology and experience with MPM images of brain that postulate a rounded lumen instead of a jittery and rough lumen, despite a very strong signal. The error example 2 illustrates a low intensity vessel junction that was judged to be an artifact by humans based on experience or information from other image planes. The error example 3 illustrates the case where a small vessel does not exhibit a strong signal and it is not connected to another major vessel.

DeepVess implements pre- and post-processing tools to deal with *in vivo* MPM images that suffer from different motion artifacts.*DeepVess* is freely available at https://github.com/mhaft/DeepVess and can be used immediately by researchers who use MPM for vasculature imaging. Also, our model can be fine-tuned further by adjusting the intensity normalization step to utilize a different part of the intensity range and training samples for other 3D vasiform structures or other imaging modalities. Similar to many machine learning solutions, *DeepVess's* performance depends on specific image features and the performance will degrade in cases where the tissues are labeled differently (e.g. vessel walls are labeled instead of blood serum) or the images intensities are concentrated in a small portion of the intensity range.

Although *in vivo* measurements present unique challenges to image segmentation, such as the red blood cell motion, in our case, we have shown that *DeepVess* successfully handles these challenges. Postmortem techniques all change the vessel diameters in the tissue processing. Hence, we believe that *in vivo* imaging is the best strategy to quantify vessel diameters. While features such as topology and length might not be affected by postmortem processing, *in vivo* imaging with MPM is important for capillary diameter measurements. Two-photon microscopy has been used to validate histology in many studies ([5, 14, 45–48]) and comparisons with other labeling techniques are quite common.

While *DeepVess* offers very high accuracy in the problem we consider, there is room for further improvement and validation, in particular in the application to other vasiform structures and modalities. For example, other types of (e.g., non-convolutional) architectures such as long short-term memory (LSTM) can be examined for this problem. Likewise, a combined approach that treats segmentation and centerline extraction methods together, such as the method proposed by Bates et al. [25] in a single complete end-to-end learning framework might achieve higher centerline accuracy levels.

Application to Alzheimer's mouse models

Capillary alteration caused by aging and Alzheimer's disease

In vivo imaging with multiphoton microscopy of capillary beds is free of distortions in vessel structure caused by postmortem tissue processing that can result in artifacts such as altered diameters [5]. However, the images often suffer from poor signal to noise and motion artifacts. An additional challenge is that unlabeled, moving red blood cells in the vessel lumen cause

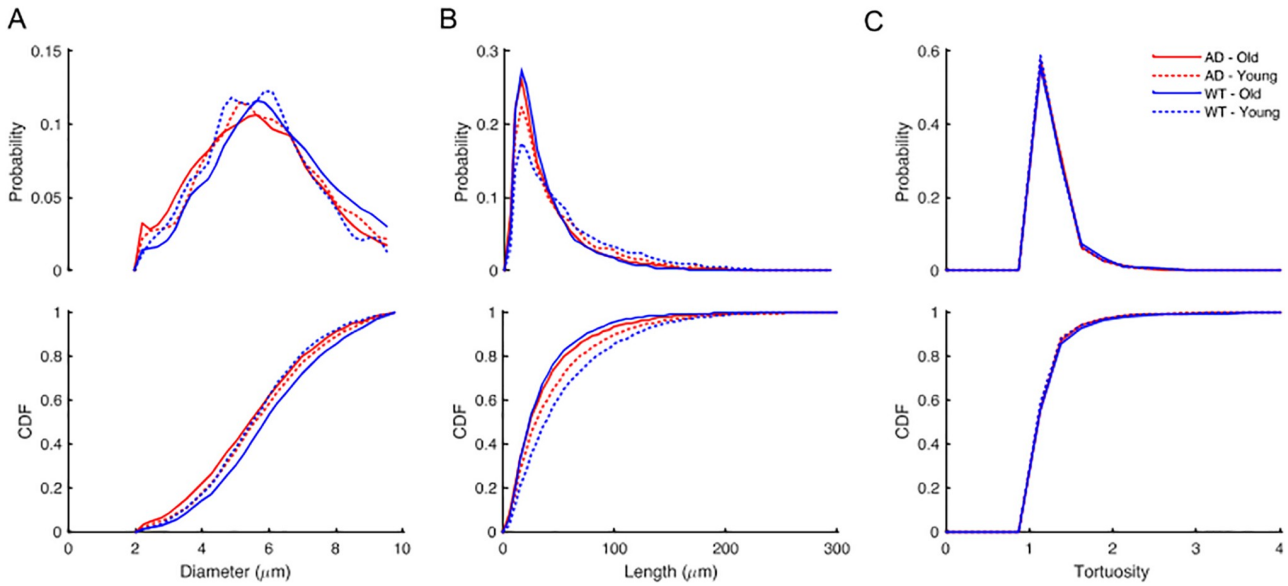

Fig 7. Comparison of capillaries between young and old mice with WT and AD genotype (6 mice in each group). The relative probability and cumulative distribution function (CDF) of the (A) diameters, (B) length, and (C) tortuosity based on all capillaries aggregated within each of the four groups. We compared these metrics between the groups using Kruskal-Wallis test followed by Bonferroni multiple comparison correction [54] (Table 2).

https://doi.org/10.1371/journal.pone.0213539.g007

dark spots and streaks that move over time. Disease models are often especially challenging because inflammation and tissue damage can further degrade imaging conditions.

Strong correlations between vascular health, brain blood flow and AD suggest that mapping the microvascular network is critical to the understanding of cognitive health in aging [49]. To explore this question, we imaged the cortical vascular networks in young and old mouse models of AD (young AD and old AD) and their young and old WT littermates (young WT and old WT). Imaged volumes ranged from 230 × 230 to 600 × 600 μm^2 in x-y and 130 to 459 μm in the z direction. We imaged 6 animals per group, with at least 3000 capillary segments analyzed for each group.

The resulting 3D stacks of images were preprocessed, segmented with *DeepVess*, and post-processed as discussed in the previous sections. Centerlines were extracted and individual vessel segments were identified. To analyze capillaries while excluding arterioles and venules, only vessel segments less than 10 μm in diameter were included [6, 50, 51]. For the vascular parameters of segment length, diameter, and tortuosity considered here, previous work has shown that AD mouse models have increased tortuosity in cortical penetrating arterioles as compared to WT mice [52, 53]. Our analysis of capillaries excluded these vessels. Three metrics were selected to characterize the vascular network. For each capillary segment, we calculated the diameter averaged along the length (Fig 7A), the length (Fig 7B), and the tortuosity, defined as the length divided by the Euclidean distance between the two ends (Fig 7C). The distributions of capillary diameter, length, and tortuosity varied little between young and old mice or between WT and AD genotype (Table 2). There were subtle shifts ($\sim 0.25\ \mu m$) in the diameter distribution between groups, but no clear differences across old/young or WT/AD and the differences in means were small compared to the standard deviation (6-27% of SD). However, we observed a decrease in the number of longer length ($>75\mu m$) capillaries in older animals as compared to young in both WT and AD mice shown by a rightward shift in the cumulative distribution function curve (Fig 7B and Table 2).

Table 2. Comparison between metrics distributions between different groups using Kruskal-Wallis test followed by Bonferroni multiple comparison correction. $\Delta\mu$ is the difference between the mean values of the two tested groups.

	Diameter (μm)		Length (μm)		Tortuosity	
	$\Delta\mu$	P-value	$\Delta\mu$	P-value	$\Delta\mu$	P-value
AD-Old vs. AD-Young	0.206	2.61E-7	7.908	7.5E-22	0.016	0.798
AD-Old vs. WT-Old	0.475	2.93E-27	2.787	0.055	0.019	0.645
AD-Old vs. WT-Young	0.095	1.20E-5	16.16	6.9E-67	0.018	0.321
AD-Young vs. WT-Old	0.269	6.39E-9	10.69	1.14E-27	0.035	0.027
AD-Young vs. WT-Young	0.110	0.012	8.252	9.12E-17	1.50E-3	1.000
WT-Old vs. WT-Young	0.379	1.1E-14	18.95	2.1E-63	0.037	0.036

https://doi.org/10.1371/journal.pone.0213539.t002

Aging and Alzheimer's disease have little effect on capillary characteristics

Using a large database of vessel segments measured in three dimensions, we surprisingly found only very small differences between groups that were dwarfed by the variance in capillary diameter or tortuosity between young and old animals or between WT and AD mouse models. The automation provided by *DeepVess* enabled the evaluation across a very large number of vessels in a large group size. The strong agreement between the measurements based on DeepVess and the manual measurements by Cruz Hernéndez et al. [6], confirms that the proposed pipeline yields unbiased and accurate metrics to analyze capillary segments. There was a decrease in the number of long capillary segments in the aged animals compared to young in both the WT and AD groups. Note that the reported metrics only represent the parietal region of cortex and that regional variability can affect our results. These finding may not generalize across all ages and mouse models of AD and could be different in other regions of the brain. Sonntag et al. [1] argue that changes in vasculature due to aging might be non-linear and multi-phasic. For instance, two studies showed that the capillary density increases during adulthood and then declines in more advanced age [55, 56]. Several previous studies have characterized the average diameters of cortical capillaries in mice, as summarized in Table 3, show high variability in results, suggesting that methodological variations make comparison between studies difficult. Other studies that compared AD models and WT also found negligible or no difference in capillary diameters. Heinzer et al. compared a different mouse model (APP23) using MRA and found no difference between WT and AD mice [57]. The same group also compared the effects of "VEGF overexpression" model and WT using SRμCT and also found little difference [58].

There are a wide range of imaging approaches used in these various studies and data from both live animal and postmortem analysis is included. It is possible that some of these differences emerge when tissues are processed rather than measured *in vivo* as was done here. Studies based on sectioned tissue sample the 3D vascular architecture differently so it is difficult to make direct comparisons between datasets. Measures of capillaries depend on the definition of capillaries. Here it was based on a threshold diameter of 10μm, which could explain some of the variability in the literature. Not surprisingly given the differences in approach and sample preparation, there is significant disagreement between reported average diameters. Some differences may, however, reflect differences in vasculature across strains and ages of animals.

Therefore, the proposed fully automated objective segmentation of 3D *in vivo* images of the vasculature can be used to reduce the variability due to sample preparation and imaging/analysis approach, allowing such strain and age differences to be elucidated clearly.

Table 3. Comparison of measured mouse capillary diameters from different studies.

Study	Background	Trans gene	Phenotype	Age (week)	Imaging Modality	VesselDiameter
This study	C57/BL6	-	WT	18-31	*in vivo* 2PEF	5.81 ± 1.62 μm
This study	C57/BL6	-	WT	50-64	*in vivo* 2PEF	6.19 ± 1.76 μm
This study	C57/BL6	APP/PS1	AD	18-31	*in vivo* 2PEF	5.92 ± 1.76 μm
This study	C57/BL6	APP/PS1	AD	50-64	*in vivo* 2PEF	5.71 ± 1.77 μm
Boero et al. [59]	BALB/C	-	WT	11	postmortem optical imaging	2.48 – 2.70 μm
Drew et al. [60]	C57/BL6	-	WT	-	*in vivo* 2PEF	2.9 ± 0.5 μm
Blinder et al. [5]	C57/BL6	-	WT	-	*in vivo* optical img., postmortem 2PEF	2 – 5.3 μm
Hall et al. [61]	C57/BL6J	NG2-DsRed	WT	-	*in vivo* 2PEF	4.4 ± 0.1 μm
Gutierrez-Jim énez et al. [51]	C57/BL6	NTac	WT	13-15	*in vivo* 2PEF	4.1 – 4.5 μm
Cudmore et al. [62]	C57/BL6	Tie2-Cre:mTmG	WT	13-21, 64, 97	*in vivo* 2PEF	5.03 ± 1.18 μm
Meyer et al. [63]	C57/BL6	APP23 & -	AD & WT	12-108	postmortem histology	4 – 6 μm
Tsai et al. [14]	Swiss	-	WT	-	*in vivo* 2PEF	3.97 – 4.11 μm
Tsai et al. [14]	C57/BL6	-	WT	-	*in vivo* 2PEF	3.97 – 4.11 μm
Heinzer et al. [57]	C57/BL6	APP23	WT	52	MRA	14 ± 5 μm
Heinzer et al. [57]	C57/BL6	APP23	AD	52	MRA	14 ± 5 μm
Heinzer et al. [64]	C57/BL6	APP23	AD	44	*SRμCT*	8.9 μm
Heinzer et al. [58]	C57/BL6	-	WT	16	*SRμCT*	5.6 ± 27.9 μm
Heinzer et al. [58]	C57/BL6	C3H/He:NSE-VEGF$_{1651}$	other	16	*SRμCT*	5.5 ± 29.3 μm
Serduc et al. [65]	Swiss nude	-	WT	5	*in vivo* 2PEF	4 – 6 μm
Vérant et al. [50]	Swiss nude	-	WT	5	*in vivo* 2PEF	8.2 ± 1.4 μm

https://doi.org/10.1371/journal.pone.0213539.t003

Conclusions

Here, we presented *DeepVess*, a 3D CNN segmentation method together with essential pre- and post-processing steps, to fully automate the vascular segmentation of 3D *in vivo* MPM images of murine brain vasculature. *DeepVess* promises to expedite biomedical research on the differences in angioarchitecture and the impact of such differences by removing the laborious, time consuming, and subjective manual segmentation task from the analysis pipelines in addition to elimination of subjective image analysis results. We hope the availability of our open source code and reported results will facilitate and motivate the adoption of this method by researchers and practitioners.

Supporting information

S1 Text. Manual 3D segmentation protocol using ImageJ.
(PDF)

S1 Fig. Jaccard as a measure of the model accuracy. The *DeepVess* results surpass the trained human annotator result at all three train, validation, and test datasets. The human annotator and *DeepVess* results are shown in dashed and solid lines respectively. The constant difference between *DeepVess* and the human annotator's results confirm the avoidance of overfitting.
(TIF)

S2 Fig. The vessel diameters measured manually in comparison to the *DeepVess's* results. There is no significant difference between two measurements (paired t-test, $n = 100$, $p = 0.34$).
(TIFF)

S1 Table. The results of investigating different field of view sizes.
(PDF)

S2 Table. The results of investigating different architectures.
(PDF)

S3 Table. The properties of six 3D images not used for training acquired from different mice included in the second independent dataset.
(PDF)

S4 Table. The results of *DeepVess* and the state-of-the-art methods on the second independent dataset from subjects not used for the model training (S3 Table). *DeepVess* surpass both of them in terms of sensitivity, Dice index, Jaccard index, and boundary modified Hausdorff distance (MHD).
(PDF)

Author Contributions

Conceptualization: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura.

Data curation: Mohammad Haft-Javaherian, Linjing Fang, Victorine Muse, Chris B. Schaffer, Nozomi Nishimura.

Formal analysis: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura.

Funding acquisition: Chris B. Schaffer, Nozomi Nishimura.

Investigation: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura.

Methodology: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

Resources: Chris B. Schaffer, Nozomi Nishimura.

Software: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura.

Supervision: Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

Validation: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

Visualization: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

Writing – original draft: Mohammad Haft-Javaherian, Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

Writing – review & editing: Chris B. Schaffer, Nozomi Nishimura, Mert R. Sabuncu.

References

1. Sonntag WE, Eckman DM, Ingraham J, Riddle DR. Regulation of cerebrovascular aging. In: Riddle DR, editor. Brain aging: models, methods, and mechanisms. Boca Raton, FL: CRC Press/Taylor & Francis; 2007. p. 279–304.
2. Hossmann K. Viability thresholds and the penumbra of focal ischemia. Annals of neurology. 1994; 36 (4):557–565. https://doi.org/10.1002/ana.410360404 PMID: 7944288
3. Lin L, Yao J, Zhang R, Chen CC, Huang CH, Li Y, et al. High-speed photoacoustic microscopy of mouse cortical microhemodynamics. Journal of biophotonics. 2017; 10(6-7):792–798. https://doi.org/10.1002/jbio.201600236 PMID: 28009098
4. Erdener ŞE, Tang J, Sajjadi A, Kılıç K, Kura S, Schaffer CB, et al. Spatio-temporal dynamics of cerebral capillary segments with stalling red blood cells. Journal of Cerebral Blood Flow & Metabolism. 2017; p. 0271678X17743877.

5. Blinder P, Tsai PS, Kaufhold JP, Knutsen PM, Suhl H, Kleinfeld D. The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. Nature neuroscience. 2013; 16 (7):889–897. https://doi.org/10.1038/nn.3426 PMID: 23749145
6. Cruz Hernández JC, Bracko O, Kersbergen CJ, Muse V, Haft-Javaherian M, Berg M, et al. Neutrophil adhesion in brain capillaries contributes to cortical blood flow decreases and impaired memory function in mouse models of Alzheimer's disease. Nature Neuroscience In press. 2019;.
7. Kalaria RN. Cerebral vessels in ageing and Alzheimer's disease. Pharmacology & therapeutics. 1996; 72(3):193–214. https://doi.org/10.1016/S0163-7258(96)00116-7
8. Farkas E, Luiten PG. Cerebral microvascular pathology in aging and Alzheimer's disease. Progress in neurobiology. 2001; 64(6):575–611. https://doi.org/10.1016/S0301-0082(00)00068-X PMID: 11311463
9. Riddle DR, Sonntag WE, Lichtenwalner RJ. Microvascular plasticity in aging. Ageing research reviews. 2003; 2(2):149–168. https://doi.org/10.1016/S1568-1637(02)00064-8 PMID: 12605958
10. Kirbas C, Quek F. A review of vessel extraction techniques and algorithms. ACM Computing Surveys (CSUR). 2004; 36(2):81–121. https://doi.org/10.1145/1031120.1031121
11. Lesage D, Angelini ED, Bloch I, Funka-Lea G. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. Medical image analysis. 2009; 13(6):819–845. https://doi.org/10.1016/j.media.2009.07.011 PMID: 19818675
12. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical Image Analysis. 2017; 42(Supplement C):60–88. https://doi.org/10.1016/j.media.2017.07.005 PMID: 28778026
13. Fraz MM, Remagnino P, Hoppe A, Uyyanonvara B, Rudnicka AR, Owen CG, et al. Blood vessel segmentation methodologies in retinal images–a survey. Computer methods and programs in biomedicine. 2012; 108(1):407–433. https://doi.org/10.1016/j.cmpb.2012.03.009 PMID: 22525589
14. Tsai PS, Kaufhold JP, Blinder P, Friedman B, Drew PJ, Karten HJ, et al. Correlations of neuronal and microvascular densities in murine cortex revealed by direct counting and colocalization of nuclei and vessels. Journal of Neuroscience. 2009; 29(46):14553–14570. https://doi.org/10.1523/JNEUROSCI.3287-09.2009 PMID: 19923289
15. Yi J, Ra JB. A locally adaptive region growing algorithm for vascular segmentation. International Journal of Imaging Systems and Technology. 2003; 13(4):208–214. https://doi.org/10.1002/ima.10059
16. Mille J, Cohen LD. Deformable tree models for 2D and 3D branching structures extraction. In: Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on. IEEE; 2009. p. 149–156.
17. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama. 2016; 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216 PMID: 27898976
18. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 2818–2826.
19. Wu A, Xu Z, Gao M, Buty M, Mollura DJ. Deep vessel tracking: A generalized probabilistic approach via deep learning. In: Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on. IEEE; 2016. p. 1363–1367.
20. Fu H, Xu Y, Wong DWK, Liu J. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on. IEEE; 2016. p. 698–701.
21. Fu H, Xu Y, Lin S, Wong DWK, Liu J. Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2016. p. 132–139.
22. Maninis KK, Pont-Tuset J, Arbeláez P, Van Gool L. Deep retinal image understanding. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2016. p. 140–148.
23. Lorthois S, Lauwers F, Cassot F. Tortuosity and other vessel attributes for arterioles and venules of the human cerebral cortex. Microvascular research. 2014; 91:99–109. https://doi.org/10.1016/j.mvr.2013.11.003 PMID: 24291593
24. Teikari P, Santos M, Poon C, Hynynen K. Deep learning convolutional networks for multiphoton microscopy vasculature segmentation. arXiv preprint arXiv:160602382. 2016;.
25. Bates R, Irving B, Markelc B, Kaeppler J, Muschel R, Grau V, et al. Extracting 3D Vascular Structures from Microscopy Images using Convolutional Recurrent Networks. arXiv preprint arXiv:170509597. 2017;.

26. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 234–241.
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; 2015. p. 3431–3440.
28. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2016. p. 424–432.
29. Jankowsky JL, Fadale DJ, Anderson J, Xu GM, Gonzales V, Jenkins NA, et al. Mutant presenilins specifically elevate the levels of the 42 residue β-amyloid peptide in vivo: evidence for augmentation of a 42-specific γ secretase. Human molecular genetics. 2003; 13(2):159–170. https://doi.org/10.1093/hmg/ddh019 PMID: 14645205
30. Denk W, Strickler JH, Webb WW, et al. Two-photon laser scanning fluorescence microscopy. Science. 1990; 248(4951):73–76. https://doi.org/10.1126/science.2321027 PMID: 2321027
31. Holtmaat A, Bonhoeffer T, Chow DK, Chuckowree J, De Paola V, Hofer SB, et al. Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. Nature protocols. 2009; 4 (8):1128–1144. https://doi.org/10.1038/nprot.2009.89 PMID: 19617885
32. Shih AY, Driscoll JD, Drew PJ, Nishimura N, Schaffer CB, Kleinfeld D. Two-photon microscopy as a tool to study blood flow and neurovascular coupling in the rodent brain. Journal of Cerebral Blood Flow & Metabolism. 2012; 32(7):1277–1309. https://doi.org/10.1038/jcbfm.2011.196
33. Pologruto TA, Sabatini BL, Svoboda K. ScanImage: flexible software for operating laser scanning microscopes. Biomedical engineering online. 2003; 2(1):13. https://doi.org/10.1186/1475-925X-2-13 PMID: 12801419
34. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophotonics international. 2004; 11(7):36–42.
35. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Medical image analysis. 1998; 2(3):243–260. https://doi.org/10.1016/S1361-8415(98)80022-4 PMID: 9873902
36. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage. 2009; 45(1):S61–S72. https://doi.org/10.1016/j.neuroimage.2008.10.040 PMID: 19041946
37. Srivastava N, Hinton GE, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. Journal of machine learning research. 2014; 15(1):1929–1958.
38. Dubuisson MP, Jain AK. A modified Hausdorff distance for object matching. In: Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on. vol. 1. IEEE; 1994. p. 566–568.
39. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Transactions on pattern analysis and machine intelligence. 1993; 15(9):850–863. https://doi.org/10.1109/34.232073
40. Kingma D, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014;.
41. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:160304467. 2016;.
42. Lee TC, Kashyap RL, Chu CN. Building skeleton models via 3-D medial surface axis thinning algorithms. CVGIP: Graphical Models and Image Processing. 1994; 56(6):462–478.
43. Jones JS, Small DM, Nishimura N. In Vivo Calcium Imaging of Cardiomyocytes in the Beating Mouse Heart With Multiphoton Microscopy. Frontiers in physiology. 2018; 9:969. https://doi.org/10.3389/fphys.2018.00969 PMID: 30108510
44. Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In: international conference on machine learning; 2016. p. 1050–1059.
45. So PT, Kim H, Kochevar IE. Two-photon deep tissue ex vivo imaging of mouse dermal and subcutaneous structures. Optics express. 1998; 3(9):339–350. https://doi.org/10.1364/OE.3.000339 PMID: 19384379
46. So PT, Dong CY, Masters BR, Berland KM. Two-photon excitation fluorescence microscopy. Annual review of biomedical engineering. 2000; 2(1):399–429. https://doi.org/10.1146/annurev.bioeng.2.1.399 PMID: 11701518
47. Tsai PS, Friedman B, Ifarraguerri AI, Thompson BD, Lev-Ram V, Schaffer CB, et al. All-optical histology using ultrashort laser pulses. Neuron. 2003; 39(1):27–41. https://doi.org/10.1016/S0896-6273(03)00370-2 PMID: 12848930

48. Zoumi A, Lu X, Kassab GS, Tromberg BJ. Imaging coronary artery microstructure using second-harmonic and two-photon fluorescence microscopy. Biophysical journal. 2004; 87(4):2778–2786. https://doi.org/10.1529/biophysj.104.042887 PMID: 15454469

49. Iadecola C. Neurovascular regulation in the normal brain and in Alzheimer's disease. Nature reviews Neuroscience. 2004; 5(5):347. https://doi.org/10.1038/nrn1387 PMID: 15100718

50. Vérant P, Serduc R, Van Der Sanden B, Rémy C, Vial JC. A direct method for measuring mouse capillary cortical blood volume using multiphoton laser scanning microscopy. Journal of Cerebral Blood Flow & Metabolism. 2007; 27(5):1072–1081. https://doi.org/10.1038/sj.jcbfm.9600415

51. Gutiérrez-Jiménez E, Cai C, Mikkelsen IK, Rasmussen PM, Angleys H, Merrild M, et al. Effect of electrical forepaw stimulation on capillary transit-time heterogeneity (CTH). Journal of Cerebral Blood Flow & Metabolism. 2016; 36(12):2072–2086. https://doi.org/10.1177/0271678X16631560

52. Dorr A, Sahota B, Chinta LV, Brown ME, Lai AY, Ma K, et al. Amyloid-β-dependent compromise of microvascular structure and function in a model of Alzheimer's disease. Brain. 2012; 135(10):3039–3050. https://doi.org/10.1093/brain/aws243 PMID: 23065792

53. Lai AY, Dorr A, Thomason LA, Koletar MM, Sled JG, Stefanovic B, et al. Venular degeneration leads to vascular dysfunction in a transgenic model of Alzheimer's disease. Brain. 2015; 138(4):1046–1058. https://doi.org/10.1093/brain/awv023 PMID: 25688079

54. Milliken GA, Johnson DE. Analysis of messy data volume 1: designed experiments. vol. 1. CRC Press; 2009.

55. Wilkinson J, Hopewell J, Reinhold H. A quantitative study of age-related changes in the vascular architecture of the rat cerebral cortex. Neuropathology and applied neurobiology. 1981; 7(6):451–462. https://doi.org/10.1111/j.1365-2990.1981.tb00245.x PMID: 7329517

56. Hunziker O, Abdel'al S, Schulz U. The aging human cerebral cortex: a stereological characterization of changes in the capillary net. Journal of gerontology. 1979; 34(3):345–350. https://doi.org/10.1093/geronj/34.3.345 PMID: 429767

57. Heinzer S, Krucker T, Stampanoni M, Abela R, Meyer EP, Schuler A, et al. Hierarchical bioimaging and quantification of vasculature in disease models using corrosion casts and microcomputed tomography. In: Proc. SPIE. vol. 5535. SPIE Bellingham, WA, USA; 2004. p. 65–76.

58. Heinzer S, Kuhn G, Krucker T, Meyer E, Ulmann-Schuler A, Stampanoni M, et al. Novel three-dimensional analysis tool for vascular trees indicates complete micro-networks, not single capillaries, as the angiogenic endpoint in mice overexpressing human VEGF 165 in the brain. Neuroimage. 2008; 39(4):1549–1558. https://doi.org/10.1016/j.neuroimage.2007.10.054 PMID: 18077185

59. Boero JA, Ascher J, Arregui A, Rovainen C, Woolsey TA. Increased brain capillaries in chronic hypoxia. Journal of applied physiology. 1999; 86(4):1211–1219. https://doi.org/10.1152/jappl.1999.86.4.1211 PMID: 10194205

60. Drew PJ, Shih AY, Kleinfeld D. Fluctuating and sensory-induced vasodynamics in rodent cortex extend arteriole capacity. Proceedings of the National Academy of Sciences. 2011; 108(20):8473–8478. https://doi.org/10.1073/pnas.1100428108

61. Hall CN, Reynell C, Gesslein B, Hamilton NB, Mishra A, Sutherland BA, et al. Capillary pericytes regulate cerebral blood flow in health and disease. Nature. 2014; 508(7494):55–60. https://doi.org/10.1038/nature13165 PMID: 24670647

62. Cudmore RH, Dougherty SE, Linden DJ. Cerebral vascular structure in the motor cortex of adult mice is stable and is not altered by voluntary exercise. Journal of Cerebral Blood Flow & Metabolism. 2016; p. 0271678X16682508.

63. Meyer EP, Ulmann-Schuler A, Staufenbiel M, Krucker T. Altered morphology and 3D architecture of brain vasculature in a mouse model for Alzheimer's disease. Proceedings of the National Academy of Sciences. 2008; 105(9):3587–3592. https://doi.org/10.1073/pnas.0709788105

64. Heinzer S, Krucker T, Stampanoni M, Abela R, Meyer EP, Schuler A, et al. Hierarchical microimaging for multiscale analysis of large vascular networks. Neuroimage. 2006; 32(2):626–636. https://doi.org/10.1016/j.neuroimage.2006.03.043 PMID: 16697665

65. Serduc R, Vérant P, Vial JC, Farion R, Rocas L, Rémy C, et al. In vivo two-photon microscopy study of short-term effects of microbeam irradiation on normal mouse brain microvasculature. International Journal of Radiation Oncology* Biology* Physics. 2006; 64(5):1519–1527. https://doi.org/10.1016/j.ijrobp.2005.11.047