Deep Learning-Based Optical-Resolution Photoacoustic Microscopy for In Vivo 3D Microvasculature Imaging and Segmentation

Huangxuan Zhao, Jia Huang, Qiang Zhou, Ningbo Chen, Liangjian Liu, Xinggang Wang, Tao Wang, Leqing Chen, Chengbo Liu,* Chuansheng Zheng,* and Fan Yang*

1. Introduction

Microcirculatory impairment significantly affects the formation of tumors and the onset of hypertension, diabetes, and various cerebrovascular diseases.[1–3] The microvasculature system helps transport nutrients to tissues and also helps remove metabolic wastes. The system is an integral part of the process of microcirculation. Important anatomical and functional information can be obtained if the microvasculature system can be imaged (in vivo) efficiently. The clinically transformed imaging techniques cannot be used to obtain high-resolution images. The resolution obtained using the computed tomography (CT) technique is not high enough. The time required to obtain high-quality images using the high-resolution magnetic resonance imaging (MRI) technique is high. The contrast achieved using the ultrasound imaging technique is significantly low, small lesions are easily overlooked. Thus, we can conclude that there is a lack of...
clinically transformable equipment that can be used to record high-quality images of the microvasculature system.

The optical-resolution photoacoustic microscopy (OR-PAM) technique is a promising microvasculature imaging technique that can be used for anatomical, functional, and molecular imaging (in situ) in real time.\cite{4,5} Endogenous and exogenous contrast agents are used during the imaging process, and the method can be used for clinical translation. The technique has been widely used in the field of preclinical research to obtain information on microvasculature-related diseases.\cite{5–7} The depth of focus (DOF) realized using OR-PAM is low and inadequate to cover the entire 3D data for microvasculature imaging. Hence, it is difficult to obtain high-quality images of the entire 3D data. When the microvessels are out of focus, the intensity of the signals is significantly reduced. This significantly hinders the process of assessment of the dynamic changes occurring in the microvasculature system during the development of microcirculation-related diseases. Thus, it is difficult to study the microvasculature system and identify the diseases using the OR-PAM technique.

Traditional image processing methods (that uses Frangi filter)\cite{8} and some advanced Frangi filter-based methods have been commonly used to detect microvessels. The Frangi filter-based methods are also used to enhance the detection ability of the methods.\cite{9–15} However, the major problem related to these methods is that the use of these methods can also result in the generation of significant extents of vascular distortions. Several image deblurring\cite{16,17} vascular enhancement\cite{18–21} and segmentation\cite{22,23} methods, based on deep learning (DL), have been proposed for photoacoustic imaging (PAI). It has been reported that the qualities of the images obtained using these methods are better than the quality of the images obtained using traditional image processing methods. The most commonly used methods are based on U-Net\cite{24} and residual dense network (RDN).\cite{25} Researchers modified the structure to achieve better results. These methods included the introduction of residual connections and dense connections. Better results were also obtained via attention mechanisms, etc. The 2D convolutional kernel is commonly used to train and test data through 2D maximum amplitude projection (MAP) or B-scan images to enhance the signal strength of the out-of-focus vessels. These methods can be used to directly recognize and extract image features following convolutional operations. Recently, the developers of U-Net have verified various biomedical data and claimed that the more the number of structural modifications (realized through the introduction of residual connections and dense connections, and via attention mechanisms), the easier it is to overfit.\cite{26} Therefore, although many modified methods can be used to realize excellent performance, the methods cannot be used universally, that is, they cannot be used to modifies all kinds of samples.

Here, we present a DL-based OR-PAM that can be used for high-quality imaging and segmentation processes with excellent generalization ability (schematic representation: Figure 1; description: 2.1. Experimental Setup). The proposed DL method is known as the Hessian matrix-based 3D context encoder network (HM-3DCE-Net). The improvement of our model is mainly due to take into consideration the characteristics of the vasculature, but not modify the model structure. The images of the microvasculature can be recorded at different depths to obtain kinds of endogenous and exogenous data. Thus, the limitations presented in the previous section can be overcome. We propose the use of the hessian matrix-based convolution (HMC) block in HM-3DCE-Net to analyze the images. This is the first deep learning network that utilizes features obtained from the second-order partial derivative of the 3D data. We chose to use the second-order derivative of 3D data because it has

![Figure 1. Schematic representation of the deep learning (DL)-based Optical-resolution photoacoustic microscopy (OR-PAM) system. ConL: convex lens; FC: fiber coupler; SMF: single-mode fiber; OBJ: objective; RM: reflection mirror; UST: ultrasonic transducer; EA: electronic amplifier; PC: personal computer; DAQ: data acquisition electronics card.](image-url)
been reported\textsuperscript{[14,27]} that the information obtained using the second-order derivative of 3D data is superior to the information obtained using the second-order derivative of 2D data. The feature map obtained using the former is better than that obtained using the latter. The intensities of the signals corresponding to the defocused microvessels were low, and it was difficult to directly identify these features using the CNN models. However, the value of the second-order partial derivative corresponding to the tubular signal in the direction perpendicular to the axis is significantly large. We use this feature to identify the microvessels using the HMC block. We combine the HMC block with decoding processes, 3D dense atrous convolution (3D-DAC) block, 3D residual multi-kernel pooling (3D-RMP) block, and decoding processes to realize the enhancement and segmentation of the vasculature. The device demonstrated excellent endogenous and exogenous vascular imaging capabilities.

\subsection*{2. Experimental Section}

\subsection*{2.1. Experimental Setup}

A custom-built OR-PAM system (reported previously\textsuperscript{[28,29]}) was upgraded for the DL-based photoacoustic microscopy-based experiments. Figure 1 presents the schematic of the new system (left) and the imaging part for samples (right, a detailed description in 2.3. Data Collection). This system primarily consists of four modules: imaging, command, trigger, and signal processing. In the imaging module, the system was equipped with an optical parametric oscillator (OPO)-pulsed laser (NT-242, Ekspla, Vilnius, Lithuania; repetition rate: 1 kHz) as the illumination source. The output of the laser beams was reshaped by an iris (Iris, ID25SS, Thorlabs, aperture size: 2 mm). The reshaped beam was focused by a condenser lens (ConvL, LA1131, Thorlabs) before being passed through a 50 $\mu$m pinhole (Pinhole, P50C, Thorlabs). Following this, it was launched into a single-mode fiber (SMF, P1-460A-FC-2, Thorlabs) coupled with the fiber coupler (FC, F-91-C1, Newport). The output of the SMF was first collimated using an objective lens (OBJ1, RMS4X, Thorlabs). Subsequently, it was reflected by a stationary reflection mirror (RM) to fill the back aperture of another identical objective lens (OBJ2), numerical aperture (NA): 0.1) to achieve optical focusing. The focused spot was illuminated on the imaging sample to generate the photoacoustic signals, which were then collected using a single-element ultrasound transducer (UST, V2022, Olympus-NDT, Kennewick, WA, USA, Center frequency: 50 MHz). The lateral and axial resolutions achieved by our system were $\approx 3.8$ and $\approx 20 \mu$m, respectively, which was measured by the edge of a sharp metallic blade. The DOF achieved by our system was 95.7 $\mu$m. In the command module, a command was given to the computer, and it was sent to the laser through the Labview software. In the trigger module, an OPO laser was used to trigger the data acquisition electronics card (DAQ) and motor controller to ensure that correct data was collected. Three axes were driven by the PI control card (Corvus PCI, Physik Instrumente, Germany). The signal processing module is the most upgraded version, the acquired signals were amplified using an electronic amplifier (EA, 5073PR, Olympus IMS). Following this, it was transmitted to the DAQ (CS1422, Gage Applied Technologies Inc., Lockport). The collected signal was processed using the Hilbert Transformation method to obtain 3D data. Following this, the data were imported into the HM-3DCE-NET.

\subsection*{2.2. Overall Description of the HM-3DCE-Net}

The overall framework of the PAI and segmentation process based on the HM-3DCE-Net is presented in Figure 2. The framework contains 2 parts: training and testing. During training, the original images were normalized to [0,1], cut into small pieces which contained 256 $\times$ 256 $\times$ 16 voxels. The 256 $\times$ 256 $\times$ 16 voxels were selected in one section depending on the computing ability of the graphics processing unit (GPU) used by us. The number of voxels in the depth direction was determined by the thickness of the imaged samples. Different samples were characterized by different thicknesses. Hence, different voxels were recorded for different depths. All data were cut into multiples of 8 (except 8) to facilitate the process of data augmentation. A total of 1280 $\times$ 256 $\times$ 256 $\times$ 16 voxel sub-images were obtained following the process of data augmentation by conducting the cropping operation (the sliding step sizes along the x/y/z directions were 128/128/8, respectively). For example, a set of 768 $\times$ 768 $\times$ 48 data could be cropped into 125 groups of sub-images following the process of data augmentation, data selection (selecting data containing background area below 95%), and imported into the HM-3DCE-Net. Of these 1280 sub-images, 204 are images of...
the iris, 402 are images of the brain, and 674 are images of the legs. The differences between the outputs and ground truths were compared to establish the superiority of the proposed method. The loss function (BCEWithLogitsLoss) was minimized, and the CNN-related parameters were continuously updated to obtain the best training model. We first used the C-scan method to obtain the original 3D data during testing. Following this, the original 3D data were normalized, cut into small pieces (256 × 256 × 16 voxels), and imported into a fully trained HM-3DCE-Net. It took an average of 0.07 s to process a small piece that contained 256 × 256 × 16 voxels. Finally, these pieces were stitched together to obtain the desired 3D results. The MAP images of the 3D results were calculated and generated. The results were presented in Results and Discussion section. The raw PAI data were used as the input data, and the binary images labeled by experienced staff were used as the appropriate labels (all training and testing (endogenous) data were labeled, and all exogenous data were kept unlabeled as it was challenging to manually label these wide-field and ultrasdense vasculature). All data were pre-processed using MATLAB (R2017a, Mathworks, Natick, MA, USA). To verify whether 1280 sub-images of data were enough for training, we used 1380, 1480, and 1580 sub-images to train models. The additional 300 sub-images were derived from the candidate training data (described in Section 2.3.1). We used these models to test data and calculate the peak signal to noise ratio (PSNR) and structural similarity index (SSIM) of the segmentation results with labels (see Appendix Table S1, Supporting Information). We found that when the number of images used for data analysis was increased to 1580, the results were not significantly affected. This proved that unless the training data increased in multiple rate, it exerted little effect on the results.

As can be seen in the HM-3DCE-Net of Figure 1, the input data sequentially passes through the HMC block, encoding process, 3D-DAC block, 3D-RMP block, and decoding processes to yield the final output. The details of the HMC block (detailed description of the HMC block has been presented in 2.2.1) are shown in Appendix Figure S1, Supporting Information. The block was specially designed to realize microvasculature imaging. The microvascular features were extracted using the second-order partial derivatives. These features were allowed to pass through trainable convolution kernels. The shapes of these convolution kernels changed with the training of the model. The changes occurred until the optimal result was obtained, and this helped correct the error attributable to second-order derivation. The distortion of images attributable to the use of the second-order partial derivatives, which is the biggest weakness of Frangi filter-based algorithms, could be avoided. It was previously stated that the microvascular features could not be extracted using the existing DL-based methods as the intensities of some of the microvascular signals were significantly low. This problem could also be addressed using the block. The encoding process consisted of three layers of Res-Net.[30] The changes in the feature map size occurring during the encoding process were presented: 16 × 256 × 256 × 1 → 16 × 256 × 256 × 32 → 8 × 128 × 128 × 64 → 4 × 64 × 64 × 128. The 3D dense atrous convolution block (3D-DAC; Appendix Figure S2, Supporting Information; detailed description has been presented in 2.2.2) and 3D-RMP (Appendix Figure S3, Supporting Information; detailed description has been presented in 2.2.3) blocks were modified using DAC and the RMP blocks,[31] and these were used between encoding and decoding. The process could be used to efficiently extract multi-scale and detailed signal features using the PAI technique. The decoding process consisted of three convolutional layers and upsampling operations. The changes in the feature map size occurring during the decoding process were presented: 4 × 64 × 64 × 132 → 8 × 128 × 128 × 128 → 16 × 256 × 256 × 128. At the end of the decoding process, a 1 × 1 × 1 convolution kernel was used to change the size of the feature map to 16 × 256 × 256 × 1. The binary cross-entropy loss was used to train the CNN model. Under conditions of inference, a softmax layer was used to generate a probability map during the process of microvascular signal enhancement. A sigmoid layer with a threshold of 0.5 was used for image segmentation. The two results share the same network and raw output logits. We implemented the network by using the PyTorch[32] library with Compute Unified Device Architecture (CUDA) and CUDA Deep Neural Network library (CUDNN) support and used the Adam optimizer[33] to minimize the loss function and update the network parameters iteratively following the process of back-propagation. The random seed was set to 0. The exponential decay learning rate[34] was used to conduct all the CNN-based experiments. The initial learning rate was set to 0.01 (decay: 0.9). The training iterations were set to 400. The models were run on an Ubuntu 16.04 LTS (64-bit) operating system equipped with an Intel Xeon E5-2680 central processing unit (CPU, 128 GB memory) and two NVIDIA GeForce GTX 1080 cards (2 × 8 GB memory). It is worth noting that the aims of using the proposed deep learning algorithm are 1) to improve the signal-to-noise ratio (SNR) of the image by enhanced photoacoustic signals; and 2) achieve image segmentation. During this whole process, the resolution did not change.

2.2.1. Detailed Description of the HMC Block

Step 1: Calculation of the second-order partial derivatives of the 3D data

\[
\begin{align*}
H_{xx}^0 &= \frac{\partial^2 I(x,y,z)}{\partial x^2} \\
H_{yy}^0 &= \frac{\partial^2 I(x,y,z)}{\partial y^2} \\
H_{zz}^0 &= \frac{\partial^2 I(x,y,z)}{\partial z^2} \\
H_{xy}^0 &= \frac{\partial^2 I(x,y,z)}{\partial x \partial y} \\
H_{xz}^0 &= \frac{\partial^2 I(x,y,z)}{\partial x \partial z} \\
H_{yz}^0 &= \frac{\partial^2 I(x,y,z)}{\partial y \partial z}
\end{align*}
\]

(1)

Here, \(I(x,y,z)\) is the input 3D data. \(H_{xx}^0, H_{yy}^0, H_{zz}^0, H_{xy}^0, H_{xz}^0, H_{yz}^0\) are the second-order partial derivatives calculated in the three directions.

Step 2: \(H_{xx}^0, H_{yy}^0, H_{zz}^0, H_{xy}^0, H_{xz}^0, H_{yz}^0\) and \(H_{zz}^0\) were fed into one convolutional layer for shallow feature extraction. One output channel was set to obtain six shallow features as the outputs.
Here, $W_{HM,1} - W_{HM}$ and $b_{HM,1} - b_{HM}$ represent the convolutional filters and biases, respectively, and $H_{1x}, H_{1y}, H_{1z}, H_{2y}, H_{2x},$ and $H_{2z}$ are the extracted shallow features.

Step 3: Merging the extracted features and original data following a $1 \times 1 \times 1$ convolution process.

### 2.2.2. Detailed Description of 3D-DAC

The 3D-DAC block is a combination of the 3D atrous convolution and Res-Net. The 3D atrous convolution is shown in Appendix Figure S2a, Supporting Information. It is obtained from the standardized 3D convolution. The atrous rate $r$ is represented in 3D. When $r = 1$ indicates a standard convolution, 3D-DAC is shown in Appendix Figure S2b, Supporting Information, and it is formed by stacking the atrous convolution under a cascade mode. The 3D-DAC considered in our study is characterized by 5 cascade branches. The first branch is solved using a conventional $3 \times 3 \times 3$ convolution and is labeled $W_{DAC,1}$. The formula to calculate the first branch is as follows.

$$F_{DAC,1} = \delta(W_{DAC,1} \times F_{encode,n} + b_{DAC,1})$$

where $W_{DAC,1}$ and $b_{DAC,1}$ are the convolutional filters and biases, respectively, and $\delta$ is the nonlinearity activation function.

$$\begin{align*}
F_{RMP,1} &= \text{upsampling}\left(\delta\left(W_{RMP,1} \times \text{pooling(branches1)} + b_{RMP,1}\right)\right) \\
F_{RMP,2} &= \text{upsampling}\left(\delta\left(W_{RMP,1} \times \text{pooling(branches2)} + b_{RMP,1}\right)\right) \\
F_{RMP,3} &= \text{upsampling}\left(\delta\left(W_{RMP,1} \times \text{pooling(branches3)} + b_{RMP,1}\right)\right) \\
F_{RMP,4} &= \text{upsampling}\left(\delta\left(W_{RMP,1} \times \text{pooling(branches4)} + b_{RMP,1}\right)\right)
\end{align*}$$

where $W_{RMP,1}, W_{RMP,2},$ and $W_{RMP,3}$ are the $1 \times 1 \times 1$ convolutions, $b_{RMP,1}, b_{RMP,2},$ and $b_{RMP,3}$ are the biases, and $\delta$ represents the nonlinear activation function.

The output obtained from the 5th branch was used as the input of the 3D-RMP, and the output of the 3D-RMP was obtained from the five branches as shown in Equation (7).

$$F_{RMP_{out}} = \text{concat}[F_{RMP,1}, \ldots, F_{RMP,4}, F_{RMP_{in}}]$$

### 2.3. Data Collection

#### 2.3.1. Endogenous PAI Data

Data from 18 healthy Balb/c mice were collected and used for training. Four mice were used for iris imaging, 5 for brain imaging, and 6 for leg imaging. Three healthy Balb/c mice (1 mouse for iris imaging, 1 mouse for brain imaging, and 1 mouse for leg imaging) were used for the process of candidate training. Another 3 healthy Balb/c mice (1 mouse for iris imaging, 1 mouse for brain imaging, and 1 mouse for leg imaging) were used for testing. The 532 nm pulsed laser of the OPO laser was used as the illumination source. For brain imaging, a surgical procedure was performed to remove the scalp. For the processes of iris and leg imaging, no pretreatment was required. During the process of imaging, the mice remained anesthetized. The mice were anesthetized using a mixture of 1.5% isoflurane gas (Euthanex, Palmer, Pennsylvania) and oxygen. Coupling gel was applied to the imaging area, and the imaging head of the photoacoustic system was placed directly above it. The fluence of the laser used was below the acceptable ANSI limit. \([16]\)
All animal handling and experimental procedures conformed to a protocol approved by the Animal Study Committee of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

2.3.2. Exogenous PAI Data

A healthy Balb/c mouse brain and a subcutaneous HepG2 tumor-bearing Balb/c mouse ear were used as the exogenous testing data. A conjugated polymer nanoparticle probe (CP NP; absorption peak: 1064 nm), the synthesis of which we have reported previously [5c], and matrigel were mixed in the ratio of 1:1. The mixture (30 μL) was injected through the tail vein before conducting the imaging process. As it is difficult to label the 3D dense network images manually, all the exogenous data were used as testing data, and the models trained with the endogenous PAI data were used to test them, and verify the generalization ability of the proposed DL model. The 1064 nm-pulsed laser of the OPO laser was used as the illumination source. The experiment was begun 5 min after injecting the CP NP. The animal handling procedures described in Section 2.3.1 were followed to conduct the experiments.

3. Results and Discussion

3.1. Overall Performance of the Proposed DL-Based OR-PAM

The test images were depth-encoded to reveal the image quality (Figure 3). Different colors in the images correspond to different depths (indicated by five color bars). The first row presents the conventional OR-PAM-based images, and the second row presents the images obtained using the DL-based OR-PAM technique. We used green to indicate the signal at the focal point in all the sub-figures. Analysis of the images reveals that as the DOF was not high, high-quality images of only a part of the blood vessels could be recorded using the conventional OR-PAM technique. The proposed new technology can be used to record high-quality images as the DL method helps increase the intensity of the microvascular signals obtained at different depths. Video S1, Supporting Information, presents the high-quality 3D images obtained for the wide-field and ultradense exogenous imaging of microvasculature systems located at different depths in the mouse brain using the proposed DL-based OR-PAM technique. Appendix Figure 4, Supporting Information, presents all second-order partial derivative images and original images. The various observations have been listed: 1) the second-order derivative images along the zz direction are very similar to the corresponding original images; we believe that the high-order features cannot be efficiently extracted as the resolution along the z-axis is low; and 2) other second-order derivative images are characterized by the presence of high-order texture features, which efficiently assist in the identification and extraction of multi-scale vasculature.

3.2. Imaging and Segmentation Results: Endogenous PAI

The well-trained RDN[35] (a 2D enhanced method that exhibits high performance in PAI[16,18]), 3D U-Net (the most popular DL method used for 3D medical imaging[24,36]), and the Frangi filter were used to process the results obtained using endogenous imaging techniques to reveal the superiority of the proposed DL-method over previously reported DL methods (Figure 4). The RDN and 3D U-Net trained under the same conditions, random seed, and hyper-parameters with the proposed method. The images of mouse iris, brain, and leg are presented in Figure 4a–c, respectively. The normalization results corresponding to endogenous PAI (for a mouse brain) were further depth encoded and presented in Video 2, Supporting Information. In each sub-image, the right side shows the enlarged view of the portion enclosed in the green dashed box outlined in the left image. The normalized results of the conventional images are shown in Figure 4a1–c1. The results obtained using the Frangi filters are shown in Figure 4a3–c3, the RDN-processed results are shown in Figure 4a5–c5, and the

Figure 3. Depth-encoded images recorded for the test samples. Scale bar = 0.5 mm.
3D U-Net-processed results are presented in Figure 4a7–c7. The results obtained using the processing technique proposed by us are shown in Figure 4a9–c9. Analysis of the enlarged images reveals that all three methods can be used to improve the quality of the images obtained using conventional OR-PAM techniques. The method proposed by us is better than the previously reported methods as small blood vessels characterized by weak signals could be efficiently imaged using this technique. A green arrow is used in each enlarged image to reveal the advantages of the proposed method over other methods. The microvessels (indicated by the green arrows) processed by the proposed method appear clearer than the microvessels processed by other methods. A varying extent of signal strength and SNR was recorded when varying depths of the 3D sample were studied using traditional OR-PAM. The results obtained using the proposed method can still be exploited to enhance the signal strength (to >0.9) of most blood vessels (the range is marked by the two blue arrows in the upper left corner of Figure 4) when the background noise is...
effectively suppressed. This proves that the proposed algorithm is efficient. This effectively helps in vessel identification, enhancement, and segmentation. However, RDN and 3D-Unet cannot be used to achieve the same effect. These can only be used to significantly enhance the signal strengths of very few blood vessels. For further verification, we randomly selected one point (yellow boxes), characterized by significantly high signal strengths (close to 1), belonging to RDN, 3D-Unet, and the method-processed image reported by us. We displayed the signal strengths of these three points on all three images. Only the intensities of the signal corresponding to these three points on the image processed by our algorithm were close to 1 (Appendix Figure S5, Supporting Information). We believe that the information corresponding to the second-order partial derivative can be efficiently obtained using our HMC module. The information could be extracted by the remaining part using our net. It is worth noting that regions where two (or more) vessels overlap appear indistinguishable in the MAP images. This cannot be attributed to the use of the algorithm developed by us, as no obvious distortion was caused by the use of the proposed method (we will further demonstrate the accuracy of the algorithm in the following sections). If the distance between the boundaries of the two blood vessels is more than the system resolution, the two blood vessels can be distinguished in the 3D image (as shown in Videos S1 and S2, Supporting Information). If the distance between the boundaries of the two blood vessels is less than the system resolution, they are labeled as one.

The segmentation results are shown in Figure 4a11–c11, and the labeled images are shown in Figure 4a13–c13. In each subimage, the right side shows the enlarged view of the image presented in the green dashed box outlined in the left image. It is difficult to detect the differences between the segmentation results with the naked eye. This proves that an excellent segmentation effect can be obtained. The signal intensity profiles were recorded using the conventional OR-PAM technique and the technique proposed by us. The images were compared to check if distortions were caused during the extraction process proposed by us. The captured images were labeled in blue, green, and red in Figure 4a2–c2, a10–c10, and a12–c12. The signal intensity profiles along the solid lines in Figure 4a2, a10, a12, b2, b10, b12, and c2, c10, c12 are shown in Figure 4a15–c15. The matching of the heights of the peak-to-peak values presented in Figure 4a15–c15 proves the accuracy and efficiency of the proposed method. Appendix Figure S6, Supporting Information, shows the errors between the labels and the segmented images reported by us. Furthermore, we randomly selected three defocused blood vessels (three yellow dotted boxes; Figure 4a1–c1). A blood vessel in the defocused area (corresponding to each yellow dotted box) is randomly selected to test the errors (Figure 4a1–c1). The signal intensity curve corresponding to these three blood vessels belonging to the original images was plotted. We calculated the maximum full width at half maximum (FWHM) of the blood vessel by analyzing the signal intensity curve of the original images. Here, FWHM represents the gold standard for vessel segmentation. We compared the value of FWHM obtained from the signal intensity curve corresponding to the segmentation results and labels (Appendix Figure S7a–c, Supporting Information) and found that both the segmentation results and labels were characterized by low-level errors.

Four key quantitative indicators (PSNR, SSIM, percent correct classification (PCC), and Dice coefficient) were used to study the similarities and differences (Table 1) between the results and labels. The mean PSNR value obtained using the method proposed by us (the enhancement and segmentation methods) is significantly high. This can be attributed to the fact that the method proposed by us can be used to enhance the signals corresponding to the blood vessels significantly, such that the signals become comparable to those of the segmentation labels. The results obtained using mean SSIM are slightly better than those obtained using U-Net and RDN, proving that these three methods hardly distort blood vessels. As we did not use U-Net and RDN to segment the images, we presented the PCC and Dice coefficient of our segmentation results compared to the labels. The mean values were 0.909 and 0.868, respectively.

### 3.3. Imaging and Segmentation Results Obtained for Wide-Field and Ultradense Exogenous PAI

The photoacoustic images of the subcutaneous HepG2 tumor-bearing mouse ear and healthy mouse brain are shown in Figure 5a,b. The conventional imaging data are shown in Figure 5a1,b1. The imaging and segmentation results obtained using the proposed DL-based OR-PAM method are shown in Figure 5a3,b3, and a5,b5, respectively. In each sub-image, the enlarged view of the region outlined by the green dashed box in the left image is shown on the right image. It is challenging to manually label the wide-field and ultradense vasculature. We did not label the images and carried out the quantitative evaluation of the system against conventional images. One small blood vessel and one low-intensity blood vessel are presented in Figure 5a,b, respectively, and were selected to determine the

| Table 1. Quantitative analysis results of peak signal to noise ratio (PSNR), structural similarity index (SSIM), percent correct classification (PCC), and Dice coefficient.                  |
|---------------------------------------------------------------|
|                  | Ins   | Brain | Leg   | Mean value |
| RDN              |       |       |       |            |
| PSNR             | 18.68 | 15.85 | 15.22 | 16.58       |
| SSIM             | 0.9818| 0.9884| 0.9483| 0.9728       |
| 3D-Unet          |       |       |       |            |
| PSNR             | 19.99 | 18.86 | 17.33 | 18.79       |
| SSIM             | 0.9701| 0.9757| 0.9336| 0.9598       |
| Our enhancement  |       |       |       |            |
| PSNR             | 26.17 | 27.29 | 24.14 | 25.87       |
| SSIM             | 0.9904| 0.9937| 0.9601| 0.9814       |
| Our segmentation |       |       |       |            |
| PSNR             | 27.61 | 28.25 | 25.01 | 26.96       |
| SSIM             | 0.9875| 0.9935| 0.9518| 0.9776       |
| PCC              | 0.929 | 0.937 | 0.861 | 0.909        |
| Dice             | 0.891 | 0.937 | 0.777 | 0.868        |
performance of the proposed method. The three curves presented in Figure 5a7, that is, blue-gray, green, and orange, represent the signal intensity of the same colored line indicated by a green arrow in Figure 5a2,a4, and a6, respectively. Similarly, the three curves in Figure 5b7 represented the signal intensity of the same colored line indicated by a green arrow in Figure 5b2, b4, and b6. The enlarged view of the region outlined by the blue dashed boxes in Figure 5a7,b7 is shown in Figure 5a8,b8, respectively. The FWHM of the Gauss-fit curve represented by the blue-gray curve was evaluated to compare the results with the results obtained using the proposed method. Although this quantitative evaluation method has limitations, it has been proved that this method can be efficiently used in the absence of labels. It can be seen from Figure 5a8,b8 that the accuracy of the proposed method (used for extracting blood vessels) was significantly high, and the differences with respect to the FWHM of the original image were within one pixel (2, 2.2 vs. 2.4, and 8, 7.9 vs. 7.3). As the high-quality results were obtained for a model trained following the process of endogenous PAI for small animals, the results presented herein for the HepG2 tumor-bearing mouse ear and healthy mouse brain demonstrated the universality of the proposed method. The results prove that the DL-based OR-PAM method proposed herein can be potentially used for various endogenous and exogenous microvasculature imaging. SNR along different depths was quantitatively assessed. The depth range corresponding to a mouse brain (depth = 1 mm) was higher than the depth range of a mouse ear (depth = 0.5 mm) and hence was selected for analysis. Appendix Figure S8a,b, Supporting Information, presents the image of the coronal projection obtained using the conventional imaging method and the
imaging method proposed by us, respectively. We calculated the SNR values at depths of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 mm (Appendix Figure S8c,d, Supporting Information) for the results obtained using the proposed method and the conventional imaging method. Quantitative analysis reveals that the SNR values of the results obtained using the proposed imaging method were better than the values obtained using the conventional method at all depths. The average improvement effect was 1.38 times.

4. Conclusion

HM-3DCE-Net was proposed to be used in the DL-based OR-PAM method after fully considering the shape characteristics of the vasculatures. It was incorporated into the custom-built OR-PAM system. The endogenous and exogenous vascular imaging results obtained while studying multiple organs were used to verify the universality and accuracy of the proposed technique. The imaging ability and segmentation ability of the method were also studied. Excellent endogenous and exogenous vasculature imaging abilities were recorded. This is the first time that high-quality imaging and segmentation results of wide-field and ultradense exogenous PAI of mouse brain vasculature located at varying depths were recorded using the OR-PAM imaging technique.

Literature[37] reviewed many exist DL PAI methods, these methods do not present the generalization ability of the proposed DL methods because they use the same part of the data for training and testing. This significantly limits the applications of the DL-based PAI equipment as the equipment will not only be used for imaging different parts but will also be used for imaging exogenous and endogenous systems. Herein, we have presented that the developed method and system exhibit a broad application prospect. The method can be used to realize endogenous and exogenous imaging of different organs. It can also be used to visualize diseased models.

Traditional PAM is either characterized by high resolution (OR-PAM) or deep penetration depth (acoustic-resolution photoacoustic microscopy, AR-PAM). These significantly limit the application range of PAM. In recent years, new methods have been proposed by updating the associated hardware or software to improve the penetration depth of OR-PAM[38] or improve the resolution of AR-PAM[39] so that several advanced PAM has both high resolution and good penetration depth. The new DL method proposed by us will work well for these devices, and further expand the application scope of PAM.

Faulty microcirculation can result in the occurrence of various diseases. It is important to develop high-performance techniques that can be used to image microvasculature systems (in vivo) to obtain anatomical, functional, and molecular information. To date, such a clinically transformed device that meets the required conditions to obtain high-quality images of the microvasculature systems has not been developed. The diagnosis and treatment of diseases related to microcirculation are realized via indirect methods.[40–42] The proposed DL-based OR-PAM methods can be efficiently used to record the images (in vivo) of the microvasculature systems. They can also be potentially used for imaging the microcirculation system in clinics.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

H.Z. and J.H. contributed equally to this work. The authors thank Prof. Lihong V. Wang for the key comments on the experiment and manuscript. The authors are grateful for financial supports from the National Natural Science Foundation of China (NSFC) (Grant nos. 82102154 and 92059108); China Postdoctoral Science Foundation (Grant nos. 2021M691156 and 2021T140238); and Youth Innovation Promotion Association CAS (Grant no. 2019352).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Keywords

deep learning, endogenous and exogenous imaging, microvasculature imaging, optical-resolution photoacoustic microscopy

Received: January 4, 2022
Revised: May 9, 2022
Published online: June 26, 2022

[1] H. A. J. Struijker-Boudier, A. E. Rosei, P. Bruneval, P. G. Camici, F. Christ, D. Hention, B. I. Levy, A. Pries, J.-L. Vanoverschelde, Eur. Heart J. 2007, 28, 2834.
[2] T. T. van Sloten, S. Sedaghat, M. R. Carnethon, L. J. Launer, C. D. A. Stethouwer, Lancet Diabetes Endocrinol. 2020, 8, 325.
[3] D. H. Wasserman, T. J. Wang, N. J. Brown, Circ. Res. 2018, 122, 1135.
[4] C. Liu, J. Chen, Y. Zhang, J. Zhu, L. Wang, Adv. Photonics 2021, 3, 016002.
[5] L. V. Wang, S. Hu, Science 2012, 335, 1458.
[6] J. Du, S. Yang, Y. Qiao, H. Lu, H. Dong, Biosens. Bioelectron. 2021, 191, 113478.
[7] S. Jeon, J. Kim, D. Lee, J. W. Baik, C. Kim, Photoacoustics 2019, 15, 100141.
[8] A. F. Frangi, W. J. Niessen, K. L. Vincken, M. A. Viergever, in International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cambridge, MA, USA 1998, pp. 130–137.
[9] L. Li, L. Zhu, C. Ma, L. Lin, J. Yao, L. Wang, K. Maslov, R. Zhang, W. Chen, J. Shi, Nat. Biomed. Eng. 2017, 1, 1.
[10] E. Merceş, J. L. Herreiz, X. L. deán-Ben, D. Razansky, Light Sci. Appl. 2019, 8, 1.
[11] B. Guo, J. Chen, N. Chen, E. Middha, S. Xu, Y. Pan, M. Wu, K. Li, C. Liu, B. Liu, Adv. Mater. 2019, 31, 1808355.
[12] L. Lin, P. Hu, J. Shi, C. M. Appleton, K. Maslov, L. Li, R. Zhang, L. V. Wang, Nat. Commun. 2018, 9, 1.
[13] L. Lin, P. Hu, X. Tong, S. Na, R. Cao, X. Yuan, D. C. Garrett, J. Shi, K. Maslov, L. V. Wang, Nat. Commun. 2021, 12, 1.
[14] A. Longo, S. Morscher, J. M. Najafabadi, D. Justel, C. Zakian, V. Ntziachristos, Photoacoustics 2020, 20, 100200.
[15] B. Lafci, E. Mercep, J. L. Herraiz, X. L. Dean-Ben, D. Razansky, Neoplasia 2020, 22, 770.
[16] H. Zhao, Z. Ke, N. Chen, S. Wang, K. Li, L. Wang, X. Gong, W. Zheng, L. Song, Z. Liu, J. Biophotonics 2019, e201960147.
[17] A. Sharma, M. Pramanik, Biomed. Opt. Express 2020, 11, 6826.
[18] H. Zhao, Z. Ke, F. Yang, K. Li, N. Chen, L. Song, C. Zheng, D. Liang, C. Liu, Adv. Sci. 2021, 8, 2003097.
[19] A. DiSpirito Iii, D. Li, T. Vu, M. Chen, D. Zhang, J. Luo, R. Horstmeyer, J. Yao, IEEE Trans. Med. Imaging 2020, 40, 562.
[20] N. Davoudi, X. L. Deán-Ben, D. Razansky, Nat. Mach. Intell. 2019, 1, 453.
[21] M. Kim, G.-S. Jeng, I. Pelivanov, M. O'Donnell, IEEE Trans. Med. Imaging 2020, 39, 3379.
[22] Y. E. Boink, S. Manohar, C. Brune, IEEE Trans. Med. Imaging 2019, 39, 129.
[23] A. Y. Yuan, Y. Gao, L. Peng, L. Zhou, J. Liu, S. Zhu, W. Song, Biomed. Opt. Express 2020, 11, 6445.
[24] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, Nat. Methods 2019, 16, 67.
[25] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, USA 2018, pp. 2472–2481.
[26] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, K. H. Maier-Hein, Nat. Methods 2021, 18, 203.
[27] H. Zhao, G. Wang, R. Lin, X. Gong, L. Song, T. Li, W. Wang, K. Zhang, X. Qian, H. Zhang, J. Biomed. Opt. 2018, 23, 046006.
[28] H. Zhao, K. Li, F. Yang, W. Zhou, N. Chen, L. Song, C. Zheng, Z. Liu, C. Liu, Opto-Electron. Adv. 2021, 4, 200017.