

A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer

Tri Vu¹ , Mucong Li¹, Hannah Humayun¹, Yuan Zhou^{1,2} and Junjie Yao¹

¹Department of Biomedical Engineering, Duke University, Durham, NC 27708, USA; ²IBM Research-China, ZPark, Beijing 100085, China
Corresponding author: Junjie Yao. Email: junjie.yao@duke.edu

Impact statement

This study has the following main impacts. It offers a promising solution for removing limited-view and limited-bandwidth artifact in PACT using a linear-array transducer and conventional image reconstruction, which have long hindered its clinical translation. Our solution shows unprecedented artifact removal ability for *in vivo* image, which may enable important applications such as imaging tumor angiogenesis and hypoxia. The study reports, for the first time, the use of an advanced deep-learning model based on stabilized generative adversarial network. Our results have demonstrated its superiority over other state-of-the-art deep-learning methods.

Abstract

With balanced spatial resolution, penetration depth, and imaging speed, photoacoustic computed tomography (PACT) is promising for clinical translation such as in breast cancer screening, functional brain imaging, and surgical guidance. Typically using a linear ultrasound (US) transducer array, PACT has great flexibility for hand-held applications. However, the linear US transducer array has a limited detection angle range and frequency bandwidth, resulting in limited-view and limited-bandwidth artifacts in the reconstructed PACT images. These artifacts significantly reduce the imaging quality. To address these issues, existing solutions often have to pay the price of system complexity, cost, and/or imaging speed. Here, we propose a deep-learning-based method that explores the Wasserstein generative adversarial network with gradient penalty (WGAN-GP) to reduce the limited-view and limited-bandwidth artifacts in PACT. Compared with existing reconstruction and convolutional neural network approach, our model has shown improvement in

imaging quality and resolution. Our results on simulation, phantom, and *in vivo* data have collectively demonstrated the feasibility of applying WGAN-GP to improve PACT's image quality without any modification to the current imaging set-up.

Keywords: Photoacoustic imaging, deep learning, artifact removal, generative adversarial network, bioimaging, photoacoustic computed tomography

Experimental Biology and Medicine 2020; 245: 597–605. DOI: 10.1177/1535370220914285

Introduction

Photoacoustic (PA) tomography (PAT) is a hybrid imaging modality combining optical excitation and ultrasonic detection. In PAT, a pulsed laser provides the excitation light that is absorbed by the biological tissue, which generates pressure waves propagating in the tissue. The pressure waves are then detected by an ultrasonic transducer or transducer array to form an image of the original optical energy deposition inside the tissue. Photoacoustic computed tomography (PACT) is a major implementation of PAT, which enables deep tissue imaging by using wide-field light illumination and parallel acoustic detection with an ultrasonic transducer array. PACT has great potential for clinical translation, due to its non-ionizing radiation, deep

penetration (>3 cm), and intrinsic functional and molecular sensitivity.^{1–3}

For most clinical applications, PACT typically uses a linear ultrasonic transducer array. Its planar detection enables flexible position on the body surface. However, it induces consequently two types of transducer-related image artifact: limited-view and limited-bandwidth artifact. The limited-view artifact in PACT is caused by incomplete signal acquisition from the partial solid angles.⁴ Due to the coherent signal generation (or specular emission) and the limited detection angle, the targets that are aligned with the transducer's acoustic axis (i.e. vertical structures) cannot be imaged.⁵ For a linear ultrasound (US) transducer array, the limited-view artifact mainly presents as curved

stripe features that stretch on both sides of the reconstructed imaged target. Similarly, the limited-bandwidth artifact is caused by the US transducer array's limited detection frequency bandwidth and acts as a band-pass filter that removes both high- and low-frequency components of the PA signals.⁶ The limited-bandwidth artifacts usually present as hollow inner features of the solid targets. In addition, the frequency bandwidth limits the axial resolution.

Various methods have been explored to reduce the reconstruction artifacts in PACT.^{4,7,8} For example, researchers have used acoustic deflectors⁸ and full-view ring-shaped transducer arrays⁹ to address the limited-view issue, which, however, substantially increased the system complexity and reduced the applicability. Various algorithmic solutions have also been studied for limited-view artifact, such as weighted-factor, iterative-based back-projection and compressed-sensing techniques.^{5,7,10–15} However, these methods are either not applicable for PACT systems with a linear-array transducer^{5,13} or time-consuming with iterative computation.^{7,10–12,14,15} To address the limited-bandwidth artifacts, Wiener filtering was used to expand the frequency spectrum of the detected signals.⁶ However, this method requires high signal-to-noise ratio (SNR) of the signal, which may be not available in practice. An iterative deconvolution method was also reported to deblur the reconstructed images, at the cost of processing time.¹⁶

Deep learning (DL) has been increasingly applied in enhancing PACT performance, including localizing wavefronts,¹⁷ improving LED-based PAT,¹⁸ and assisting cancer detection.^{19–21} DL has also been extensively explored for PACT artifact removal. For example, several groups have reported the use of UNet and other deep convolutional neural networks (CNNs) to address the limited-view and sparse-sampling issues as postprocessing correction,^{22–24} direct reconstruction,²⁵ and model-based learning.²⁶ For the band-limited response, fully connected neural network has also been applied to the radio-frequency (RF) data to retrieve out-of-band signals.²⁷ Similar studies by Allman *et al.*²⁸ focused on locating true point targets in the presence of reflection artifacts. All the methods have shown promising results on simulated and experimental data; however, none have addressed both limited-view and limited-bandwidth artifacts simultaneously for PACT systems with a linear-array transducer.

Recent advances in DL for medical imaging include one of the most important breakthroughs: generative adversarial networks (GANs). By employing two models competing with each other, GAN is known for generating realistic synthetic images from arbitrary input.²⁹ More details on GAN will be discussed in the 'Materials and methods' section. Applications of GAN in medical imaging include image reconstruction and segmentation, as well as disease diagnosis in X-ray CT and MRI.^{30,31} The downside of GAN is the unstable training. It often suffers from vanishing gradient and mode collapse when either of the models becomes too strong for the other one.³² Different techniques have been studied to stabilize GAN, including noise addition³³ and

loss modification,^{34,35} which have shown improved performance and outcome.^{36,37}

Here, we propose a deep-learning-based solution to address PACT's limited-view and limited-bandwidth artifacts. Our method is based on stabilized Wasserstein generative adversarial network with gradient penalty (WGAN-GP) and additional mean-squared-error (MSE) loss function. WGAN-GP combines UNet³⁸ and DCGAN²⁹ to provide artifact-reduced PACT images with significantly improved quality. WGAN-GP was trained and tested on simulated data generated by the k-Wave toolbox,³⁹ with the time-reversal reconstruction images as the model input. The network's *in vivo* performance was validated on mouse vasculature images, and compared with time-reversal and UNet-based results. Overall, WGAN-GP has shown superior performance in removing limited-view and limited-bandwidth artifacts in PACT images.

Materials and methods

WGAN – GP and model architecture

A typical GAN model has two CNNs: the generator (G) and discriminator (D). The two parts compete in a min-max problem described in equation (1)⁴⁰

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim P_d(x)} [\log D(x)] + \mathbb{E}_{z \sim P_z(z)} \left(\log \left(1 - D(G(z)) \right) \right) \quad (1)$$

In this loss function, G takes the time-reversal reconstructed image with artifacts as the input z , and outputs the model-resolved image $G(z)$ that can 'trick' D to classify $G(z)$ as an artifact-free image. P_d and P_z are the real and the artifact-heavy data distribution respectively. D , on the other hand, is trained to assign the correct label to $G(z)$ as a model-resolved output and x as an artifact-free image (ground truth). In this way, G is trained to produce the best possible output in order to make D misclassify $G(z)$, while D is trained to be more sensitive to identify G 's attempt.

The loss function (1) in vanilla GAN suffers from training instability due to the vanishing gradient of D when minimizing the Jensen–Shannon divergence between P_d and P_z . Alternatively, in WGAN, Wasserstein metric, or Earth's Mover distance, $W(d, z)$ is proposed. $W(d, z)$ is differentiable almost everywhere under mild assumptions, thus leading to a more stable optimization of G . In this case, D becomes an optimal discrimination as a critic, instead of a classifier. This critic is under the 1-Lipschitz constraint enforced by a gradient penalty which ensures the gradient norm to be 1. Last but not the least, an additional MSE loss component is added to maintain the information of the reconstructed images. Yang *et al.*³⁷ suggested the use of perceptual loss based on pretrained VGG network instead of MSE. However, since the reconstructed PACT images are not similar to the ImageNet data employed for the pretrained VGG model,⁴¹ we will first use the MSE loss in

this study. Overall, our final loss function is

$$\min_G \max_D V_{GAN}(G, D) = -\mathbb{E}_x [D(x)] + \mathbb{E}_z \left[D(G(z)) \right] + \lambda_1 GP(\hat{x}) + \lambda_2 MSE(G(z), x) \quad (2)$$

With : $GP(\hat{x}) = E_{\hat{x}} \left[\left(\|\nabla \hat{x} D(\hat{x})\|_2 - 1 \right)^2 \right]$

$MSE(G(z), x) = \mathbb{E}_{(x, z)} \left[\frac{1}{N^2} \|G(z) - x\|_2^2 \right]$

In this equation, $\mathbb{E}(\cdot)$ denotes the expectation operator, N is the total number of pixels, and \hat{x} is sampled from $G(z)$ and x with t uniformly sampled between 0 and 1: $\hat{x} = tG(z) + (1 - t)x$. The weighting parameter λ_1 of the gradient penalty takes the suggested value of 10,³⁴ while λ_2 is 20 based on our training experiments. The number of critic iterations per generator iteration is 5.

The overall architecture of WGAN-GP is shown in Figure 1. G in WGAN-GP takes the form of UNet³⁸ layers. In Figure 1, each block of the UNet shows its name (DnL: down-sampling layer; UL: up-sampling layer) and the number of channels. The breakdown of each step in the dashed box shows its associated layers with their corresponding parameters. UNet has been proven effective in

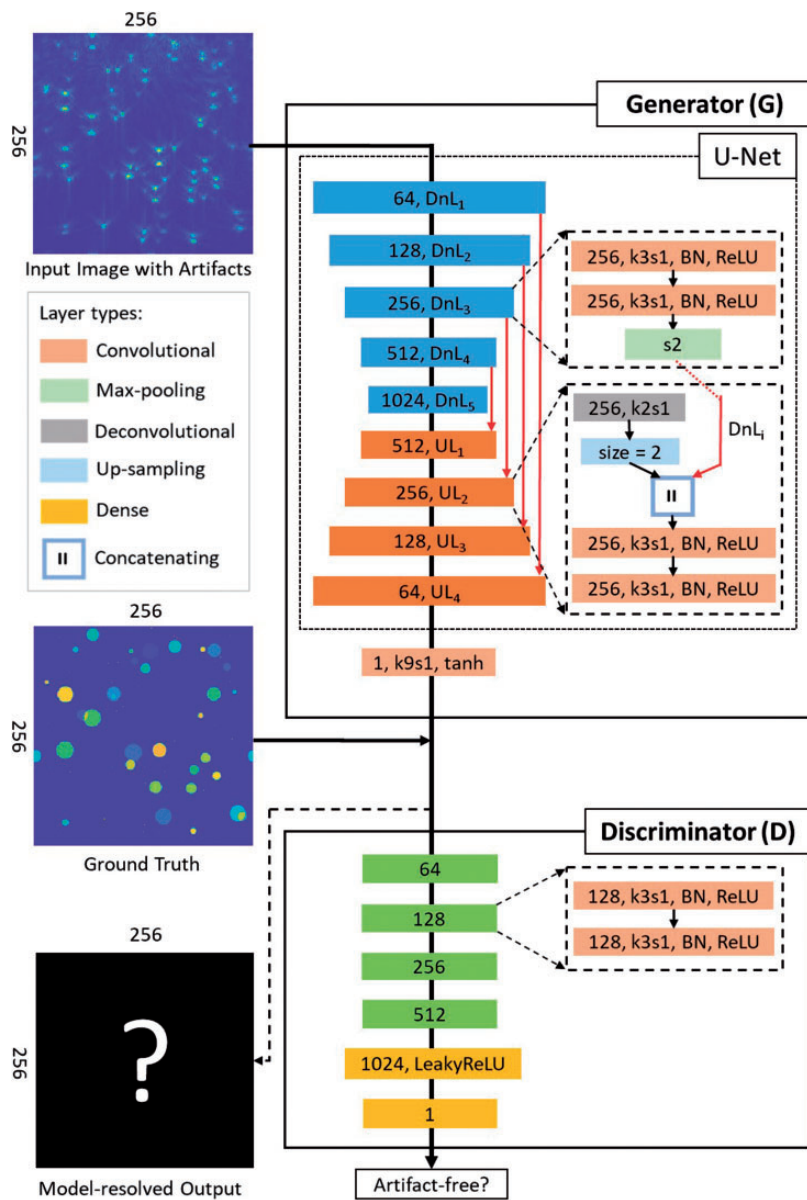


Figure 1. WGAN-GP model architecture. In all layers, the first number is the number of filters. k and s denote the kernel size and stride, respectively. In up-sampling layer, $size$ is the upsampling factor. (A color version of this figure is available in the online journal.)
 BN: batch normalization; ReLU: rectified linear unit.

passing features from lower layers and achieving faster training without re-learning redundant features. Critic D follows the model construction proposed by DCGAN with enlarging feature maps toward the output, enabling an increasing number of learned features for accurate critic of the input. Different from the conventional GAN discriminator, D in WGAN-GP does not have the sigmoid layer for the final output.

Data-sets, simulation, and model training

In addition to building the actual model, constructing the training and testing data-set is a vital step for WGAN-GP. Our study used the k-Wave toolbox to simulate the training and testing data in PACT.³⁹ To evaluate the performance of WGAN-GP, we simulated a data-set with 5000 images of randomly position-distributed disks (4000 for training and 1000 for testing). These disks have diameters ranging from 0.15 to 3 mm. To train the model with *in vivo* data, we adapted the brain vascular database by two-photon microscopy (TPM), as shown in Figure 3(e).⁴² There are three reasons that make TPM vascular data useful for training WGAN-GP. First of all, blood vessels are the major targets in many PACT applications. TPM images, as shown in Figure 3(e), contain cross-sectional blood vessels that closely ensemble the targets in PACT images. Secondly, the TPM images have much better resolutions than the simulated PACT images, and thus can serve as the ground truth. Thirdly, the TPM images have high-density high-variance vascular structures, which help train the models to learn *in vivo* conditions with complex vascular network. For each stack of TPM volumetric images, we removed the first 50 and the last 30 depth stacks which either have oversized blood vessels or significantly lower SNR. All the input TPM images for k-wave simulation were rotated 90° in order to emphasize the longitudinal vessels. To determine the training data size, a common practice is to obtain the learning curve which shows the testing accuracy versus the training size. For DL, it is intuitive that we need as many data as possible to improve the performance of the model. The learning curves from other studies have shown that the accuracy of the model increases logarithmically with training data size.⁴³ Thus, for disk data, we empirically determined the training size based on our targeted training time of 16 hours for WGAN-GP model. For TPM vascular data, because the total data have ~9000 samples, a train/test split of 80/20 gives us 7200 training and 1800 testing instances.

The training and testing data generated from simulation were then carefully crafted to be close to the realistic data. In our simulation, each disk or TPM vascular image with the size of 256 by 256 pixels was passed through the k-Wave toolbox to generate the raw RF photoacoustic data and the reconstructed image with limited-view and limited-bandwidth artifacts. White noise was added to the RF data before reconstruction with randomly assigned SNR ranging from 10 to 100. All the reconstructed data were normalized between 0 and 1 after Hilbert transform. Throughout the entire simulation, we follow the configuration of L7-4 linear transducer array, which has 128 elements, a central frequency of 5 MHz, and a 3 dB detection

Table 1. Training time for UNet and WGAN-GP with disk and TPM vascular data.

Training time (hours)	UNet	WGAN-GP
Disk data	2.7	16.0
TPM vascular data	6.0	29.2

TPM: two-photon microscopy; WGAN-GP: Wasserstein generative adversarial network with gradient penalty.

bandwidth of 60%. The size of the reconstructed image size was 38.4 mm by 38.4 mm with a grid size of 150 μm . The medium in the simulation was tissue with an average speed of sound of 1540 m/s. Directivity of the transducer was also taken into account based on the receiving aperture of each element.

Both WGAN-GP and UNet models were trained on disk and TPM vascular data. All the training for WGAN-GP was performed on an NVIDIA RTX 2080 Ti GPU using Keras with Tensorflow back-end. The number of training epochs was 50 with a batch size of five for each iteration in each epoch. The training time for each model is summarized in Table 1. It is not surprising that WGAN-GP takes more time to train than UNet, because of increased model complexity and multiple D updates in one iteration. For a fair comparison between UNet and WGAN-GP, we allowed UNet model training with both data-sets to converge, as shown in Figure 2. All the codes of the models with complete parameters can be found at <https://github.com/trivu169/2d-artifact-removal-PACT>. All the data used in the paper can be shared upon request.

Evaluation metrics and experimental data

The evaluation matrices for the model with simulated data include the structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR). SSIM and PSNR were used to compare the similarity between a reference image x and a target image $G(z)$. SSIM contains local information that reflects the interdependence of neighboring pixels and is relatively perceptual.⁴⁴ PSNR, on the other hand, resides more on global information owing to its dependence on MSE between two images. SSIM and PSNR were calculated for each output of WGAN-GP and then averaged over all the testing data.

Experimental data were then employed to evaluate the model's performance. The DL models trained with disk data-set were tested on two types of phantom data. The first phantom was a group of transparent plastic tubes (diameter: 1.5 mm) filled with black ink to provide optical absorption. The tubes were placed at different depths (5 mm step size) in an optically scattering medium with a reduced scattering coefficient of $\sim 7 \text{ cm}^{-1}$ at 1064 nm. This phantom set-up mimicked the simulated disk data and evaluated the sensitivity of the models to light attenuation and SNR. The second phantom was a point target represented by a human hair with a diameter of $\sim 100 \mu\text{m}$ embedded in clear agar. Band-limited frequency response of the linear transducer array resulted in a blurred point-spread function (PSF). To demonstrate WGAN-GP's performance

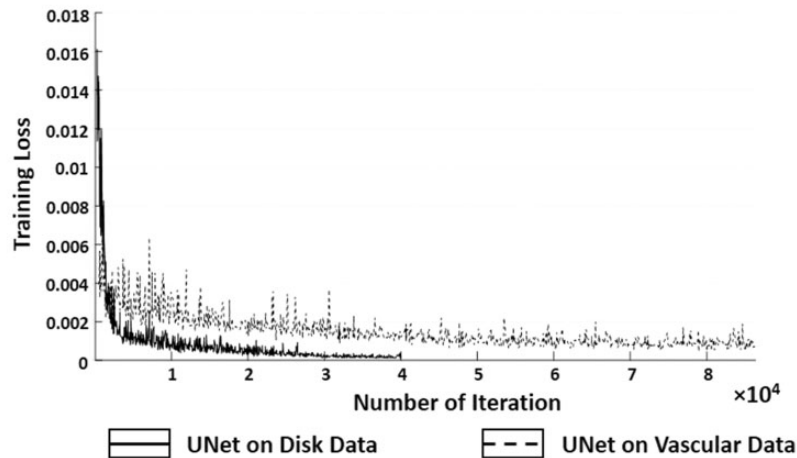


Figure 2. Training loss of UNet over iterations.

to reduce limited-bandwidth artifact, the model should be able to improve the PSF to a narrower profile.

For *in vivo* data, we imaged the skin vasculature in the trunk of a female mouse. The protocol was approved by the Institutional Animal Care and Use Committee of Duke University. The *in vivo* data were used to evaluate the DL models trained with TPM vascular data-set. For all experimental data, we used our PACT system based on a commercial US scanner (Vantage 128, Verasonics) with an L7-4 linear transducer array and an Nd:YAG laser (Q-smart 850, Quantel).⁴⁵ Laser pulses at 532 nm were used, with a pulse energy of 30 mJ and repetition rate of 10 Hz. The experimental set-up is shown in Figure 6(a).

Results

Simulated testing data

The performance of WGAN-GP compared with time-reversal and UNet on the simulated disk data is shown in Figure 3. For the simulated disk data, both WGAN-GP and UNet have improved the reconstructed time-reversal images that have heavy limited-view artifacts (i.e. the curved stripe features) and limited-bandwidth artifacts (i.e. the hollow disks). It is clear that the DL-resolved images can recover the solid inner structures of the disks, which are largely missing in the time-reversal images, illustrated by the structures in the red boxes (Figure 3(a) to (d)). WGAN-GP and UNet have on par performance on simulated disk data, represented by the SSIM and PSNR results in Table 2. For TPM vascular data, WGAN-GP shows a slightly better performance than UNet. Besides higher SSIM and PSNR, WGAN-GP outperforms UNet in recovering low-contrast structures, shown by the yellow arrows in Figure 3(e) to (h). As the PA sources become more complex, WGAN-GP shows superior performance over UNet. Additionally, as shown by the white dashed arrows in Figure 3(e) to (h) and close-up images in Figure 3(i) to (l), both models can better recover vertical structures than the time-reversal result.

Phantom data

For the tube phantom, both DL models trained with disk data-set have better performance than time-reversal. At the first four depths, similar to the results on simulated disks, WGAN-GP and UNet recover the inner structures of the tube, as shown in Figure 4(a) to (c). Due to light attenuation, the SNR of the blood tube decreases with depth (Figure 4(a) and (e)). We can observe that when the SNR is less than ~ 15 dB, both UNet and WGAN-GP cannot reconstruct the true target, as illustrated in Figure 4(b) and (c). Nonetheless, for *in vivo* imaging, PACT can provide sufficient SNR in deep tissues (17.5 dB at ~ 4 cm⁴⁶), even with compact laser sources.⁴⁷ In addition, our model has great denoising capability. As shown in Figure 4(e), WGAN-GP improves the SNR by 3 and 1.5 folds over time-reversal and UNet, respectively.

The capability of the models to resolve a point target was then evaluated by the hair phantom data. The PSF of the time-reversal result clearly shows the enlarged axial width due to band-limited frequency response, and the wing-shaped limited-view artifacts (Figure 5(a)). The DL model results have improved the PSFs as illustrated in Figure 5(b) and (c). The spatial profiles shown in Figure 5(d) further show that WGAN-GP has better spatial resolution than UNet. The difference is more significant in the lateral direction, where WGAN-GP's and UNet's FWHM are 40 and 55 μ m respectively. Additionally, the stripe artifacts due to limited view (red arrows in Figure 5(a)) are largely removed by WGAN-GP, but not by UNet.

In vivo data

Finally, the performance of WGAN-GP on *in vivo* mouse vascular data is shown in Figure 6, with a representative cross-sectional PACT image. Compared to the time-reversal images (Figure 6(c)), the WGAN-GP images provide substantially improved visibility for continuous vascular structures, including the vertical vessels (Figure 6(e) and (g)). UNet has comparable improvement as shown in Figure 6 (d) and (f); however, UNet recovers fewer details than WGAN-GP as denoted by the close-up images in

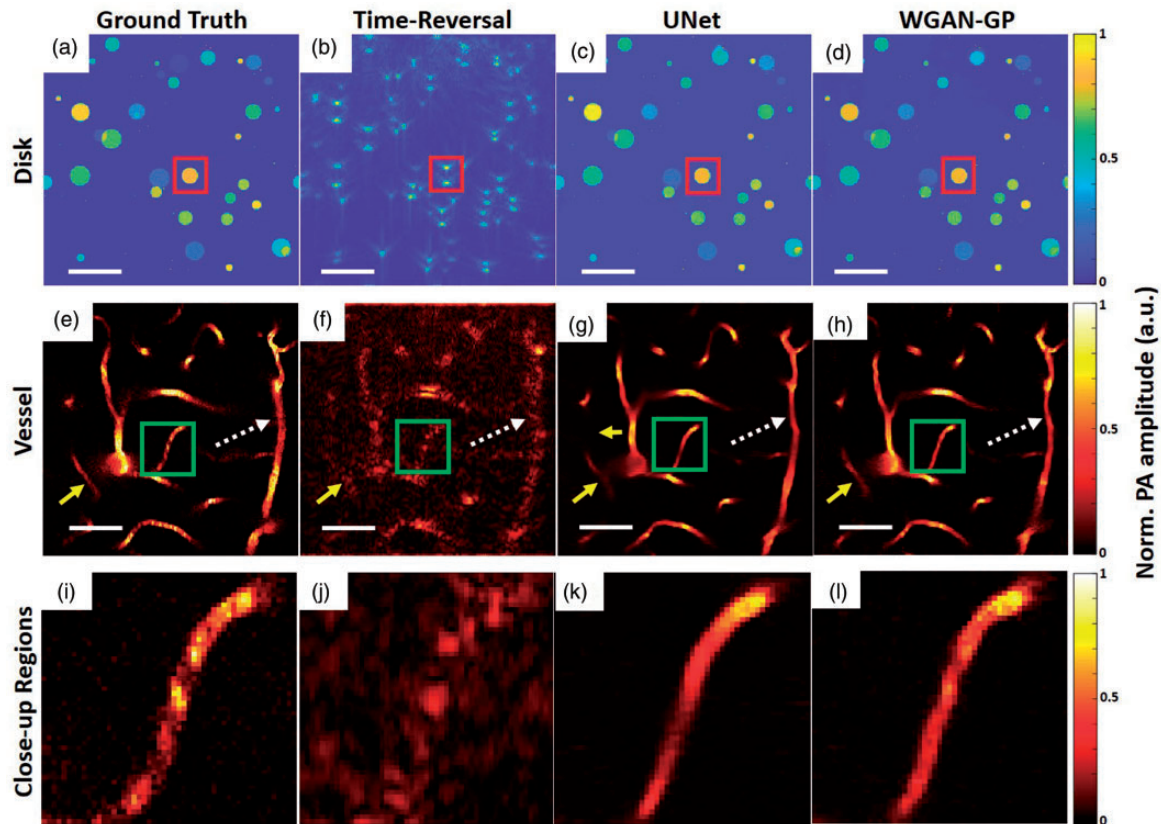


Figure 3. Performance of WGAN-GP on simulated disk and TPM vascular images. Representative ground truth, time-reversal reconstruction, UNet and WGAN-GP results on (a–d) disk and (e–h) TPM vascular images. (i–l) Close-up images of the green boxes in (e–h) showing the vertical structures. The red boxes in (a–d) highlight the superior performance of DL models to recover the true disk shapes. The yellow arrows in (e–h) highlight the low-contrast structures that are reconstructed by WGAN-GP only. The dashed white arrows point out the vertical vessels recovered by the models. Scale bar: 5 mm. (A color version of this figure is available in the online journal.)

Figure 6(f) and (g). Moreover, WGAN-GP also has better contrast-to-noise ratio (CNR) of 9.3 than UNet (CNR = 6.1). We believe that the CNR improvement is due to the difference in the loss function between the two models. UNet uses the MSE for gradient descent, which leads to more image blurry and thus lower CNR. This blurring effect is also evident in the simulated results (Figure 3(i) to (l)), where UNet cannot reconstruct the fine details. For *in vivo* data, in which the PA targets are dense (Figure 6(c)), UNet's blurring effect is more significant, resulting in lower CNR.

Discussion

In this work, we have reported a new deep-learning model WGAN-GP, which is based on the GAN, to reduce the limited-view and limited-bandwidth artifacts in PACT images. We first trained and tested WGAN-GP using simulated disk data and TPM vascular images, showing improved SSIM and PSNR. WGAN-GP then shows superior performance on the experimental phantom and *in vivo* animal data compared to UNet both qualitatively and quantitatively. Overall, WGAN-GP is capable of reducing the limited-view and limited-bandwidth artifacts of PACT using a linear transducer array, and thus, has great potential to enhance the image quality without modifying the imaging system or reducing the imaging speed. Such

Table 2. SSIM and PSNR results of UNet and WGAN-GP on disk and vessel testing data.

	UNet	WGAN-GP
SSIM (mean \pm standard deviation)		
Disk	0.96 \pm 0.03	0.96 \pm 0.03
Vessel	0.62 \pm 0.09	0.65 \pm 0.08
PSNR (mean \pm standard deviation)		
Disk	32.2 \pm 3.14	32.1 \pm 3.14
Vessel	25.7 \pm 2.22	26.5 \pm 2.06

PSNR: peak signal-to-noise ratio; SSIM: structural similarity index; WGAN-GP: Wasserstein generative adversarial network with gradient penalty.

improvements by WGAN-GP are beneficial for a variety of PACT applications, such as mapping the tumor vasculature in thermal ablation and detecting blood clots during sonothrombolysis.

Nonetheless, our method has some limitations. First, in our forward model, we only consider the US generation but not the optical excitation. In practice, different light delivery strategies lead to different optical fluence distribution within the sample. Our model has not been trained to accommodate the optical fluence variation in the training data. In the next step, we will incorporate optical excitation in our simulation as an end-to-end forward model. Secondly, the DL networks are target-specified and can only recognize the PA targets similar to those used in the

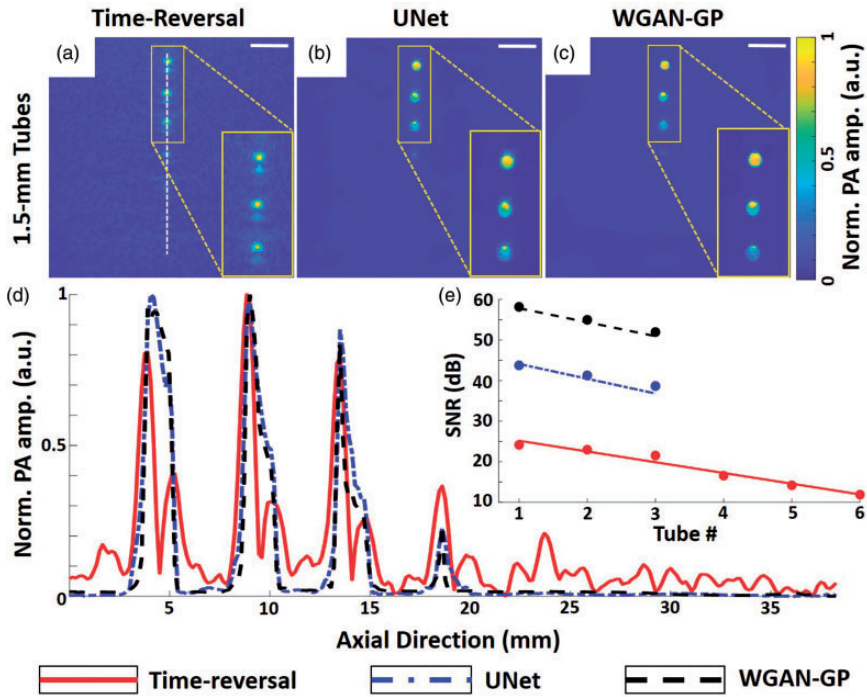


Figure 4. Experimental performance of WGAN-GP on multiple-depth tube phantom. (a–c) Reconstructed images by time-reversal, UNet, and WGAN-GP, respectively. (d) Axial profiles along the stacked tubes, indicated by the dashed white line in (a). (e) Corresponding SNRs at each depth for different method. Scale bar: 5 mm. (A color version of this figure is available in the online journal.) PA: photoacoustic; SNR: signal-to-noise ratio; WGAN-GP: Wasserstein generative adversarial network with gradient penalty.

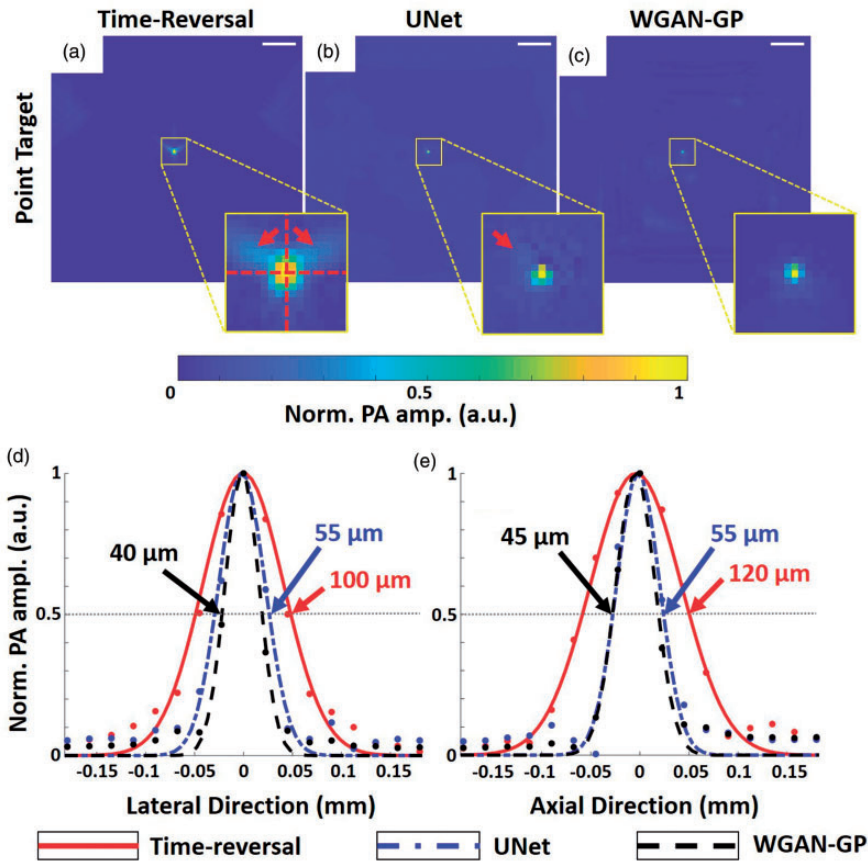


Figure 5. Experimental performance of WGAN-GP on a hair phantom. (a–c) Results from time-reversal, UNet, and WGAN-GP, respectively. Red arrows denote the stripe artifacts due to limited view. (d and e) Lateral and axial profiles along the dashed red lines in (a). Scale bar: 5 mm. (A color version of this figure is available in the online journal.) PA: photoacoustic; WGAN-GP: Wasserstein generative adversarial network with gradient penalty.

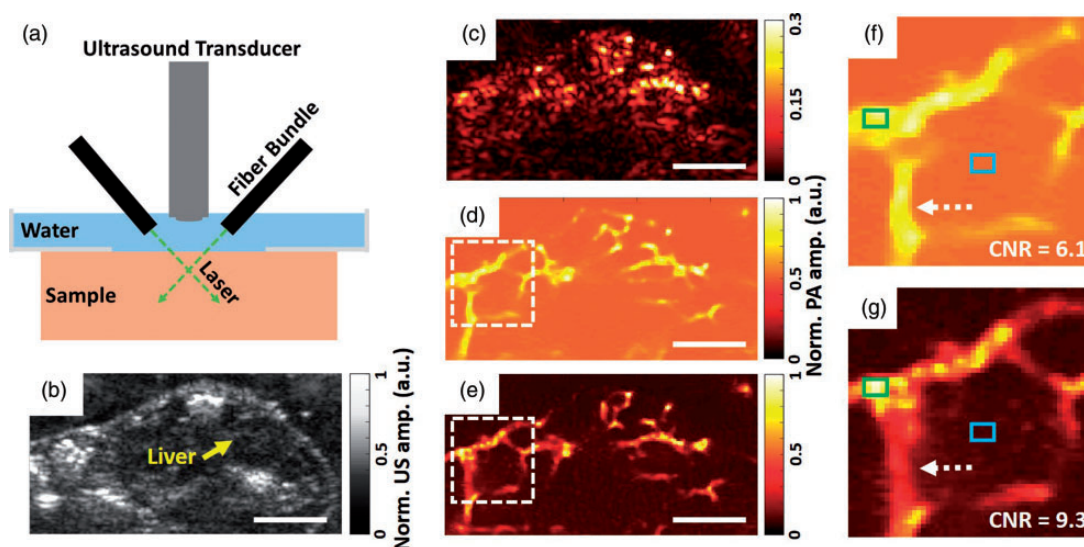


Figure 6. Experimental performance of WGAN-GP on *in vivo* trunk vascular images of a mouse. (a) System set-up for the experimental data. (b) Cross-sectional B-mode US image of the mouse trunk with the corresponded PACT images from (c) time-reversal, (d) UNet, and (e) WGAN-GP. (f and g) Close-up images of the region indicated by the white dashed boxes in UNet and WGAN-GP images, respectively. For CNR calculation, the green boxes denote the target, while the blue boxes indicate the background regions. The dashed white arrows highlight the vertical vessels recovered by the models. Scale bar: 7.5 mm. (A color version of this figure is available in the online journal.)

CNR: contrast-to-noise ratio; PA: photoacoustic; US: ultrasound.

training data, which is a common issue for DL. Therefore, the network trained with disk data cannot be applied for *in vivo* vascular data. Future work will create more generic training data with various types of PA targets. Finally, the training time of WGAN-GP is two times longer than that of UNet. It will be even more time-consuming for 3D image reconstruction.

Our future work will mainly focus on optimizing the model accuracy and validating the *in vivo* results. Currently, the SSIM of the model on the vessel data is only about 0.65, which can be improved by using a larger training data size and incorporating other vascular databases as PACT ground truth. In addition, even though the model has been thoroughly validated by the experimental phantom data with simple targets (tubes and hair), it remains technically challenging to validate the fidelity of the recovered structures for the *in vivo* images. This is the common problem for current DL approaches on reconstruction enhancement. US imaging can be used as a concurrent validation. However, conventional B-mode US imaging has limited sensitivity to blood vessels. In the future, we propose to use PACT-compatible microbubbles to enhance the US imaging of blood vessels and thus validate our DL model.

Authors' contributions: All the authors contributed to the experimental and model design, data analysis, and manuscript writing.

ACKNOWLEDGMENT

We thank Yuqi Tang for editing the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: National Institute of Health (1R01EB028143, R01 NS111039, R21 EB027304, R43 CA243822, R43 CA239830, R44 HL138185); Duke MEDx Basic Science Grant; Duke Center for Genomic and Computational Biology Faculty Research Grant; Duke Institute of Brain Science Incubator Award; American Heart Association Collaborative Sciences Award (18CSA34080277).

ORCID iD

Tri Vu  <https://orcid.org/0000-0002-1384-8647>

REFERENCES

- Kim J, Park S, Jung Y, Chang S, Park J, Zhang Y, Lovell JF, Kim C. Programmable real-time clinical photoacoustic and ultrasound imaging system. *Sci Rep* 2016;**6**:35137
- Diot G, Metz S, Noske A, Liapis E, Schroeder B, Ovsepian SV, Meier R, Rummeny E, Ntziachristos V. Multispectral optoacoustic tomography (MSOT) of human breast cancer. *Clin Cancer Res* 2017;**23**:6912–22
- Valluru KS, Willmann JK. Clinical photoacoustic imaging of cancer. *Ultrasonography* 2016;**35**:267–80
- Xu Y, Wang LV, Ambartsoumian G, Kuchment P. Reconstructions in limited-view thermoacoustic tomography. *Med Phys* 2004;**31**:724–33
- Burgholzer P, Bauer-Marschallinger J, Grün H. Weight factors for limited angle photoacoustic tomography. *Phys Med Biol* 2009;**54**:3303–14
- Cao M, Feng T, Yuan J, Xu G, Wang X, Carson PL. Spread spectrum photoacoustic tomography with image optimization. *IEEE Trans Biomed Circuits Syst* 2017;**11**:411–9

7. Liu X, Peng D, Ma X, Guo W, Liu Z, Han D, Yang X, Tian J. Limited-view photoacoustic imaging based on an iterative adaptive weighted filtered backprojection approach. *Appl Opt* 2013;**52**:3477–83
8. Huang B, Xia J, Maslov K, Wang L. Improving limited-view photoacoustic tomography with an acoustic reflector. *J Biomed Opt* 2013;**18**:110505
9. Xia J, Chatni M, Maslov K, Guo Z, Wang K, Anastasio M, Wang L. Whole-body ring-shaped confocal photoacoustic computed tomography of small animals in vivo. *J Biomed Opt* 2012;**17**:050506
10. Haltmeier M, Sandbichler M, Berer T, Bauer-Marschallinger J, Burgholzer P, Nguyen L. A sparsification and reconstruction strategy for compressed sensing photoacoustic tomography. *J Acoust Soc Am* 2018;**143**:3838–48
11. Sandbichler M, Krahmer F, Berer T, Burgholzer P, Haltmeier M. A novel compressed sensing scheme for photoacoustic tomography. *SIAM J Appl Math* 2015;**75**:2475–94
12. Provost J, Lesage F. The application of compressed sensing for photoacoustic tomography. *IEEE Trans Med Imaging* 2008;**28**:585–94
13. Burgholzer P, Bauer-Marschallinger J, Grün H. Experimental evaluation of reconstruction algorithms for limited view photoacoustic tomography with line detectors. *Inverse Probl* 2007;**23**:S81–S94
14. Ma S, Yang S, Guo H. Limited-view photoacoustic imaging based on linear-array detection and filtered mean-backprojection-iterative reconstruction. *J Appl Phys* 2009;**106**:123104
15. Haltmeier M, Berer T, Moon S, Burgholzer P. Compressed sensing and sparsity in photoacoustic tomography. *J Opt* 2016;**18**:114004
16. Rejesh NA, Pullagurla H, Pramanik M. Deconvolution-based deblurring of reconstructed images in photoacoustic/thermoacoustic tomography. *J Opt Soc Am A Opt Image Sci Vis* 2013;**30**:1994–2001
17. Johnstonbaugh K, Agrawal S, Durairaj DA, Fadden C, Dangi A, Karri SPK, Kothapalli S-R. A deep learning approach to photoacoustic wavefront localization in deep-tissue medium. *IEEE Trans Ultrason Ferroelectr Freq Control* 2020. Epub ahead of print. DOI: 10.1109/TUFFC.2020.2964698.
18. Anas EMA, Zhang HK, Kang J, Boctor EM. Towards a fast and safe LED-based photoacoustic imaging using deep convolutional neural network. In: *International conference on medical image computing and computer-assisted intervention*, 2018, pp.159–67. Berlin: Springer
19. Jnawali K, Chinni B, Dogra V, Rao N. Automatic cancer tissue detection using multispectral photoacoustic imaging. *Int J Cars* 2019;**15**:309–320
20. Zhang J, Chen B, Zhou M, Lan H, Gao F. Photoacoustic image classification and segmentation of breast cancer: a feasibility study. *IEEE Access* 2018;**7**:5457–66
21. Rajanna AR, Ptucha R, Sinha S, Chinni B, Dogra V, Rao NA. Prostate cancer detection using photoacoustic imaging and deep learning. *Electron Imaging* 2016;**2016**:1–6
22. Antholzer S, Haltmeier M, Schwab J. Deep learning for photoacoustic tomography from sparse data. *Inverse Probl Sci Eng* 2019;**27**:987–1005
23. Guan S, Khan A, Sikdar S, Chitnis P. Fully dense UNet for 2D sparse photoacoustic tomography artifact removal. *IEEE J Biomed Health Inform* 2019;**24**:568–76
24. Davoudi N, Deán-Ben XL, Razansky D. Deep learning photoacoustic tomography with sparse data. *Nat Mach Intell* 2019;**1**:453–60
25. Guan S, Khan A, Sikdar S, Chitnis PV. Pixel-wise deep learning for improving image reconstruction in photoacoustic tomography. *J Acoust Soc Am* 2019;**145**:1811–1811
26. Hauptmann A, Lucka F, Betcke M, Huynh N, Adler J, Cox B, Beard P, Ourselin S, Arridge S. Model-based learning for accelerated, limited-view 3-d photoacoustic tomography. *IEEE Trans Med Imaging* 2018;**37**:1382–93
27. Gutta S, Kadimesetty VS, Kalva SK, Pramanik M, Ganapathy S, Yalavarthy PK. Deep neural network-based bandwidth enhancement of photoacoustic data. *J Biomed Opt* 2017;**22**:116001
28. Allman D, Reiter A, Bell M. Photoacoustic source detection and reflection artifact removal enabled by deep learning. *IEEE Trans Med Imaging* 2018;**37**:1464–77
29. Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken AP, Tejani A, Totz J, Wang Z. Photo-realistic single image super-resolution using a generative adversarial network. In: *CVPR2017*, IEEE, p.4
30. Yi X, Babyn P. Sharpness-aware low-dose CT denoising using conditional generative adversarial network. *J Digit Imaging* 2018;**31**:655–69
31. Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Guo Y. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans Med Imaging* 2017;**37**:1310–21
32. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Analysis* 2019;**58**:101552
33. Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: an overview. *IEEE Signal Process Mag* 2018;**35**:53–65
34. Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville AC. Improved training of Wasserstein GANs. In: *Advances in neural information processing systems*, NIPS, 2017, pp.5767–77
35. Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: *International conference on machine learning*, PMLR, 2017, pp.214–23
36. Quan TM, Nguyen-Duc T, Jeong W-K. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Trans Med Imaging* 2018;**37**:1488–97
37. Yang Q, Yan P, Zhang Y, Yu H, Shi Y, Mou X, Kalra MK, Zhang Y, Sun L, Wang G. Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Trans Med Imaging* 2018;**37**:1348–57
38. 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*, 2015, pp.234–41. Berlin: Springer
39. Treeby BE, Cox BT. k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. *J Biomed Opt* 2010;**15**:021314
40. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: *Advances in neural information processing systems*, NIPS, 2014, pp.2672–80
41. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M. Imagenet large scale visual recognition challenge. *Int J Comput Vis* 2015;**115**:211–52
42. Uhlirou H, Tian P, Kılıç K, Thunemann M, Sridhar VB, Bartsch H, Dale AM, Devor A, Saisan PA. Neurovascular Network Explorer 2.0: a database of 2-photon single-vessel diameter measurements from mouse SI cortex in response to optogenetic stimulation. *Front Neuroinform* 2017;**11**:4
43. Sala V. Power law scaling of test error versus number of training images for deep convolutional neural networks. In: *Multimodal sensing: technologies and applications*. Bellingham WA: International Society for Optics and Photonics, SPIE, 2019, p.1105914
44. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 2004;**13**:600–12
45. Li M, Lan B, Sankin G, Zhou Y, Liu W, Xia J, Wang D, Trahey G, Zhong P, Yao J. Simultaneous photoacoustic imaging and cavitation mapping in shockwave lithotripsy. *IEEE Trans Med Imaging* 2019;**39**:468–477
46. Kim C, Erpelding TN, Jankovic L, Pashley MD, Wang LV. Deeply penetrating in vivo photoacoustic imaging using a clinical ultrasound array system. *Biomed Opt Express* 2010;**1**:278–84
47. Wang D, Wang Y, Wang W, Luo D, Chitgupi U, Geng J, Zhou Y, Wang L, Lovell JF, Xia J. Deep tissue photoacoustic computed tomography with a fast and compact laser system. *Biomed Opt Express* 2017;**8**:112–23