

Deep learning for image reconstruction in thermoacoustic tomography

Qiwen Xu(徐启文), Zhu Zheng(郑铸), and Huabei Jiang(蒋华北)

Citation: Chin. Phys. B, 2022, 31 (**2**): 024302. DOI: 10.1088/1674-1056/ac0dab
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Thermoacoustic assessment of hematocrit changes in human forearms

Xue Wang(王雪), Rui Zhao(赵芮), Yi-Tong Peng(彭亦童), Zi-Hui Chi(迟子惠), Zhu Zheng(郑铸), En Li(李恩), Lin Huang(黄林), and Hua-Bei Jiang(蒋华北)

Chin. Phys. B, 2021, 30 (**9**): 094302. DOI: 10.1088/1674-1056/ac041c

High-contrast imaging based on wavefront shaping to improve low signal-to-noise ratio photoacoustic signals using superpixel method

Xinjing Lv(吕新晶), Xinyu Xu(徐新羽), Qi Feng(冯祺), Bin Zhang(张彬), Yingchun Ding(丁迎春), Qiang Liu(柳强)

Chin. Phys. B, 2020, 29 (**3**): 034301. DOI: 10.1088/1674-1056/ab6842

Underwater acoustic metamaterial based on double Dirac cone characteristics in rectangular phononic crystals

Dong-Liang Pei(裴东亮), Tao Yang(杨洮), Meng Chen(陈猛), Heng Jiang(姜恒)

Chin. Phys. B, 2019, 28 (**12**): 124301. DOI: 10.1088/1674-1056/ab55d2

Influence exerted by bone-containing target body on thermoacoustic imaging with current injection

Yan-Hong Li(李艳红), Guo-Qiang Liu(刘国强), Jia-Xiang Song(宋佳祥), Hui Xia(夏慧)

Chin. Phys. B, 2019, 28 (**4**): 044302. DOI: 10.1088/1674-1056/28/4/044302

Acoustic beam splitting in a sonic crystal around a directional band gap

Ahmet Cicek, Olgun Adem Kaya, Bulent Ulug

Chin. Phys. B, 2013, 22 (**11**): 114301. DOI: 10.1088/1674-1056/22/11/114301

Deep learning for image reconstruction in thermoacoustic tomography

Qiwen Xu(徐启文)[1], Zhu Zheng(郑铸)[2], and Huabei Jiang(蒋华北)[3,†]

[1] *School of Electronic Science and Engineering (National Exemplary School of Microelectronics), University of Electronic Science and Technology of China, Chengdu* 611731, *China*
[2] *School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing* 400065, *China*
[3] *Department of Medical Engineering, University of South Florida, Tampa, FL* 33620, *USA*

(Received 30 April 2021; revised manuscript received 11 June 2021; accepted manuscript online 23 June 2021)

Microwave-induced thermoacoustic tomography (TAT) is a rapidly-developing noninvasive imaging technique that integrates the advantages of microwave imaging and ultrasound imaging. While an image reconstruction algorithm is critical for the TAT, current reconstruction methods often creates significant artifacts and are computationally costly. In this work, we propose a deep learning-based end-to-end image reconstruction method to achieve the direct reconstruction from the sinogram data to the initial pressure density image. We design a new network architecture TAT-Net to transfer the sinogram domain to the image domain with high accuracy. For the scenarios where realistic training data are scarce or unavailable, we use the finite element method (FEM) to generate synthetic data where the domain gap between the synthetic and realistic data is resolved through the signal processing method. The TAT-Net trained with synthetic data is evaluated through both simulations and phantom experiments and achieves competitive performance in artifact removal and robustness. Compared with other state-of-the-art reconstruction methods, the TAT-Net method can reduce the root mean square error to 0.0143, and increase the structure similarity and peak signal-to-noise ratio to 0.988 and 38.64, respectively. The results obtained indicate that the TAT-Net has great potential applications in improving image reconstruction quality and fast quantitative reconstruction.

Keywords: thermoacoustic tomography (TAT), TAT reconstruction, deep learning, finite-element

PACS: 43.35.Ud, 87.85.Pq, 87.57.nf, 84.35.+i **DOI:** 10.1088/1674-1056/ac0dab

1. Introduction

Microwave-induced thermoacoustic tomography (TAT) is an emerging hybrid imaging technique based on tissue electromagnetic properties, and provides high microwave contrast, high spatial resolution, and deep penetration depth imaging.[1–8] It has so far been widely applied to the detection of breast,[9–12] kidney,[13] and prostate[14] cancers as well as angiography[15,16] and brain imaging.[17] In the TAT, the target absorbs a microwave electromagnetic pulse, producing an ultrasonic wave due to thermal expansion.[18] The ultrasonic wave propagates out and is received by an ultrasonic transducer, and the inverse problem can be solved by an image reconstruction algorithm such as delay-and-sum (DAS),[19–21] time reversal,[22–24] and model-based iterative methods.[25–31] The DAS is a most commonly used algorithm in TAT which can achieve fast reconstruction, but suffers streaking artifacts, resulting in image distortion, especially in the situation of limited view or sparse view.[32] The model-based iterative algorithms can obtain better imaging results and achieve quantitative reconstruction in terms of tissue conductivity, but its iterative process is time consuming.

Deep learning has achieved great success in the field of computer vision and has been widely used and developed in the field of medical imaging. In recent years, several image reconstruction methods based on deep learning have been proposed, covering computed tomography (CT),[33–40] magnetic resonance imaging (MRI),[41–43] photoacoustic tomography (PAT),[44–49] and other fields.[50] Inspired by the success in these fields, we introduce deep learning method into TAT image reconstruction. Since this paper is the first study on deep learning in the field of TAT, and the reconstruction principles of PAT and TAT are similar, we take PAT for example to analyze the existing methods. The deep learning-based reconstruction method can be divided mainly into post-processing method and direct method. The post-processing reconstruction is to use the convolutional neural network (CNN) to post-process the low quality initial pressure images produced from a traditional reconstruction method, which has been shown to well solve these problems such as limited/sparse view, and denoising. However, the reconstruction error from the traditional reconstruction methods will give rise to further significant artifacts and missing of image features. The direct reconstruction method leverages the information from the sensor data to provide end-to-end image reconstruction, and the work for PAT reconstruction using numerical simulations has demonstrated the feasibility of this method. Waibel *et al.* used

†Corresponding author. E-mail: hjiang1@usf.edu
© 2022 Chinese Physical Society and IOP Publishing Ltd

a modified U-Net to estimate the initial pressure distribution directly from the PA raw sensor data.[51] Anas *et al.* used a dense convolutional neural networks to achieve photoacoustic beamforming.[52] Lan *et al.* proposed DU-Net with auxiliary loss for multi-frequency ring-shape transducer array.[53] However, since the method requires the model to learn the transition between the signal domain and image domain, it requires a great number of data and stronger model representation. Owing to these limitations, the existing direct reconstruction method does not perform so well as the post-processing method. In addition to the above two methods, several innovative ideas were recently proposed. For instance, a hybrid method uses two encoders to extract features from the sensor data and beam formed images, respectively, and reconstruct initial energy density with combined features.[54,55] A model-based method uses a prior constraints learned by training a CNN for image reconstruction.[56,57]

To the best of our knowledge, this is the first time that deep learning has been used for image reconstruction in TAT. In this work, we explore the reasons for the inaccurate reconstruction with standard U-Net model and propose a new network called TAT-Net to achieve end-to-end reconstruction from sinogram to image in TAT with high quantitative accuracy. In addition, facing the scarcity of TAT experimental data and the lack of training set, a new paradigm for data generation using finite element method (FEM) is implemented, and a signal processing method is combined to solve the problem of domain gap between synthetic data and realistic experimental data. The TAT-Net trained with synthetic data is evaluated in both simulations and phantom experiments to demonstrate the superiority of the proposed method over the traditional reconstruction methods and the potential applications in the situation of experimental data scarcity.

2. Methods

In the TAT, the acoustic inversion from the measured data to an estimate of the initial acoustic pressure distribution is an ill-posed problem due to the incomplete or imperfect measurement data and inaccurate forward model. For the image reconstruction tasks in TAT, we hope to find a reconstruction operator A to achieve the mapping form measurement data S to initial acoustic pressure P. The deep learning can help to solve this problem by representing the reconstruction operator A with a CNN which includes a set of learnable parameters θ, such as $P = A_\theta(S)$ where θ can be determined by a suitable optimization method with the given training data.

2.1. Deep learning for image reconstruction

The neural network used for image reconstruction is constructed based on the encoder–decoder structure. The corresponding neural network architecture of the TAT-Net is illustrated in Fig. 1. In the TAT-Net, the encoder part is responsible for the image downsampling, which continuously extract the image features to obtain the advanced feature maps. Conversely, the decoder part implements continuous up-sampling, which maps the extracted features to the final reconstructed image. The input of the TAT-Net is the two-dimensional (2D) sinogram which is processed by arranging the sensor data collected in all directions (*i.e.*, sinogram in a size of 270×180 represents that the number of channels is 180 and the data length of each channel is 270), and the output is the initial energy density distribution. The input image is fed into a 7×7 convolution layer with 2×3 stride to obtain the square feature map. The encoder part consists of 12 blocks, including 8 B1 blocks and 4 B2 blocks, and the decoder part also consists of 12 blocks, including 7 B1 blocks and 5 B3 blocks. The pooling layer and up-sampling operator are replaced by convolution layer and transposed convolution layer with a stride of 2×2 for higher representation capacity. To avoid the checkerboard artifacts,[58] the size of the transposed convolution kernel is set to be twice the stride.

Unlike previous neural networks for reconstruction task such as U-Net, the global residual learning and feature fusion between encoder and decoder are not recommended in TAT-Net. There exists a significant difference between the input sinogram and output reconstructed image, and learning the residual between them is more difficult, as using global residual in this task leads to a high probability of unconvergent training. The shallow feature fusion will result in shallow features being output directly into the final image without sufficient convolution for feature translation, which introduces some sinogram-like artifacts into the reconstructed image (see Subsection 3.2). In addition, we observe porphyritic BN artifacts[59] in the experiments. Thus, we remove all batch normalization layers from residual blocks for stable training and consistent performance.

Since we remove the batch normalization layers, training a deep network becomes difficult. We adopt the residual module[60] to solve the problem of gradient disappearance. To reduce training time and memory consumption, bottleneck structure is used in the entire TAT-Net. The B1 block is a standard bottleneck residual block. In the bottleneck residual structure CNN, traditional up-sampling and down-sampling operations take up a large part of the overall time and computation of the network, so we design the downsampling blocks B2 and up-sampling blocks B3 with bottleneck residual structure. The residual shortcut used in the down-sampling block is a 2×2 stride convolution and up-sampling is 2×2 stride transposed convolution.

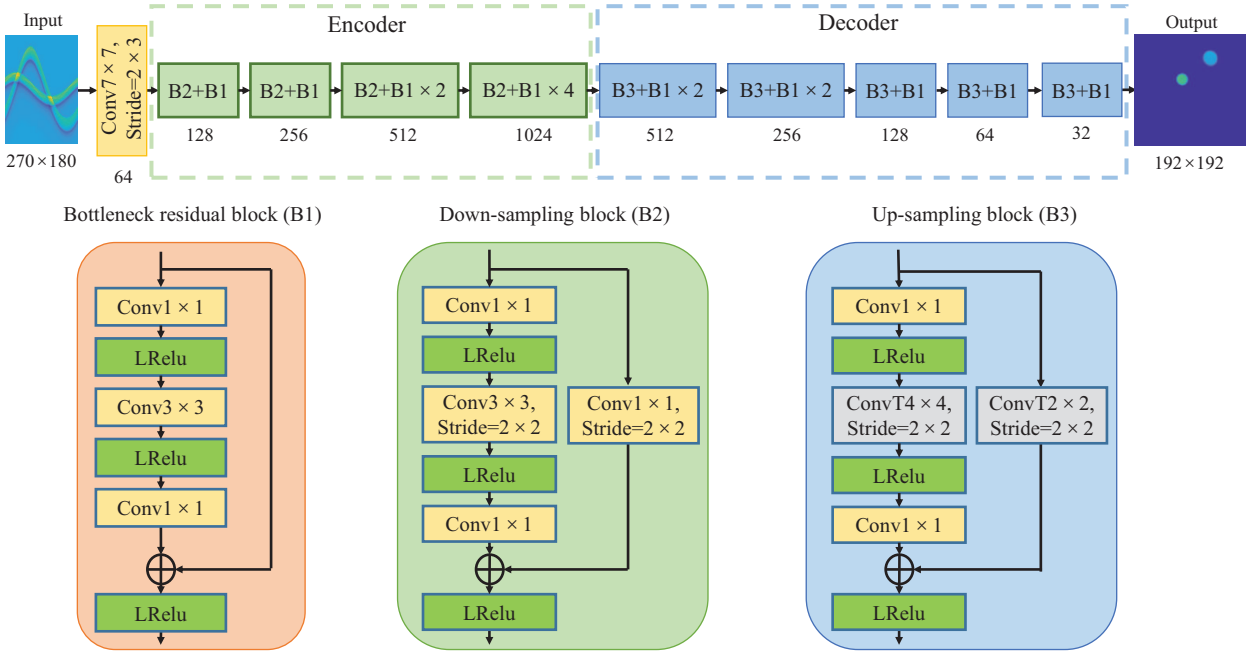


Fig. 1. Detailed network architecture of TAT-Net for image reconstruction, where the plus and multiplication signs in the colored rectangle represent the connection and repetition times of the block, respectively, the numbers written below the rectangle indicate the channels of feature map, the stride size of each convolution layer is 1 by default, except for those already indicated, and ConvT indicates the transposed convolution layer.

The TAT-Net uses leaky rectified linear units ($a = 0.01$) as activation function to obtain a steady training. We initialize the weights in convolution layers and transpose the convolution layer into a Kaiming normal distribution ($a = 0.01$).[61] The parameters of the network are optimized by minimizing the mean square error (MSE) loss function:

$$L = \|GT - f(X)\|_2^2, \tag{1}$$

where GT is the ground truth, X is the sinogram image, and f is TAT-Net operator.

We continuously reduce the MSE loss between the model output and GT by Adam algorithm[62] to optimize the parameters of the model. The learning rate and batch size set to be 10^{-4} and 8, respectively. In the prediction process, we input the sinograms into the trained model to obtain the initial energy loss density images.

In the experiments, the network is implemented in PyTorch framework.[63] and deployed in a computer platform with Intel(R) Xeon(R) CPU E5-2630 v4 2.2 GHz and NVIDIA TITAN Xp with 12G memory.

2.2. Synthetic data generation

The deep learning-based method depends on a large quantity of data; however, the current experimental conditions of TAT do not have the capacity for data mass production. In addition, the data and images acquired through the experiments can be affected by various factors, such as system instabilities, difference among the reconstruction methods, and inhomogeneous distribution of the microwave field, *etc*. These issues will impede us from obtaining realistic target parameters, and thus, we seek for a reliable simulation method to construct the synthetic dataset that satisfies the mapping of the real initial energy density distribution into the sensor data.

Our data generation method includes two steps. The first step is to solve the forward solution of the thermoacoustic wave equation through the finite element method (FEM),[64] and the thermoacoustic wave equation is defined as

$$\nabla^2 p(\boldsymbol{r},t) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} p(\boldsymbol{r},t) = -\frac{\beta \varphi(\boldsymbol{r})}{C_\mathrm{p}} \frac{\partial J(t)}{\partial t}, \tag{2}$$

where c is the acoustic speed, β is the volume expansion coefficient, C_p is the specific heat, $\varphi(\boldsymbol{r})$ denotes the energy deposition in the tissue, and $J(t)$ represents the microwave pulse. Equation (2) is based on the assumption that microwave power dissipation in a region is so rapid that thermal diffusion can be neglected. $J(t) = \delta(t - t_0)$ is also assumed in our study.

In the FEM simulation, a total of 180 acoustic receivers are equally distributed along the surface of the 80-mm diameter circular background region. The finite-element mesh used for the forward calculations has 5371 nodes and 10560 elements. One to three circles of different sizes and energy densities are deployed as initial energy densities, and these targets are randomly placed at different positions in the region of interest (ROI) with 40 mm×40 mm in size. We assume that the distribution of ultrasonic velocity and microwave energy are uniform in simulation. We can obtain sensor data $S^c(N)$ from each acoustic receiver's channel, where N is the data length.

The second step is to preprocess the data. In order to reduce the computational cost of the neural network, we reduce the input size of the network by deleting redundant data. According to the size of ROI and the location number of the acoustic receivers, we can calculate a data range that con-

tributes significantly to the reconstructed image:

$$S^c_{\text{new}} = S^c(i), \quad i \in \left(\frac{f_s(R_0 - a)}{v}, \frac{f_s(R_0 + a)}{v} \right), \tag{3}$$

where f_s is the sampling frequency, R_0 is the radius of the circular acoustic receiver array, a is the size of square ROI, and v is the ultrasonic velocity. We use the following parameter configuration, $f_s = 10$ MHz, and $v = 1495$ m/s. In addition, Gaussian noise and random bias are added into the sinogram to improve the ability to generalize the model and to reduce the domain gap between synthetic and realistic data. The processed sinogram has a signal-to-noise ratio (SNR) of 20 dB.

Finally, we use 270×180 pixel sinogram as the network input and 192×192 pixel energy density image as the ground truth. The dataset consists of 8192 training sets and 1024 validation sets. In order to verify the robustness of the model in terms of the number of imaging targets, an additional data containing 4 or 5 circular targets are generated.

2.3. TAT phantom data acquisition

To further demonstrate the feasibility of our proposed method, we conduct several phantom experiments with a TAT system shown in Fig. 2(a) to obtain the phantom experiment data. Plastic tubes containing physiological saline solution are used as the reconstruction target, and the photograph of the three tubes is shown in Fig. 2(b). In the TAT system, a microwave pulse is generated by a custom-designed 3-GHz microwave generator (bandwidth: 50 MHz, peak power: 70 kW, pulse duration: 750 ns), couples into a rectangular waveguide and is irradiated to a phantom by a horn antenna (114 mm×144 mm). The averaged microwave power density at the phantom surface is less than 0.32 mW/mm^2, which is far below the safety standard (20 mW/cm^2 at 3 GHz). A single-element unfocused transducer (V323, Olympus, central frequency: 2.25 MHz, nominal bandwidth: 1.35 MHz) is used to acquire the TA signal over 360° at 180 positions around the sample rotated by the step motor. The digital signals are amplified by an amplifier and then recorded by a data acquisition card (PCI4732, Vidts Dynamic, sampling frequency: 50 MHz). The data are down sampled from 50 MHz to 10 MHz and the ROI is the same as that in the simulation setup. In the experiment, the scanning radius of the probe is 164 mm. In order to make the experimental data and simulation data have the same scanning radius property, a method of processing the signal is used to ensure the same detector positions as the simulation setup by setting the virtual detectors.[65] As with the processing of synthetic data, we perform data refinement on the phantom data according to Eq. (3). Since the reflected signals created by non-target components in a TAT system factor are often present in these deleted data, the data refinement has the additional benefit of avoiding misinterpretation of the network due to the reflected signals. So far, we have obtained 270×180 pixel thermoacoustic experimental sinogram as the test dataset.

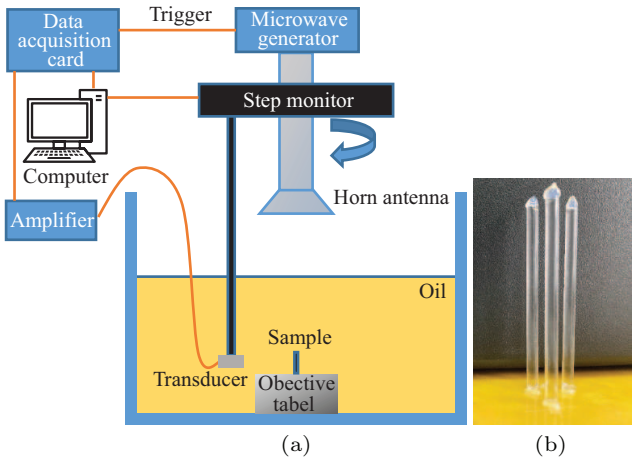


Fig. 2. (a) Schematic diagram of experimental setup, and (b) photograph of three saline-containing tubes.

3. Experiments and results

3.1. Synthetic data

We evaluate the performance of the proposed method by comparing the reconstruction results of DAS, U-Net, and TAT-Net. Figure 3 shows the results from the synthetic data with 1 to 3 circular targets. We can find that the TAT-Net can accurately reconstruct the position and shape of each target. Comparing with DAS, the image from TAT-Net has more well-defined boundaries and purer background. The U-Net provides more satisfactory results than DAS but produces additional artifacts in the 3-circle case. To visually show the reconstruction of TAT-Net, in Fig. 4 are plotted the one-dimensional (1D) line intensity profiles passing through the red dashed lines shown in Fig. 3 for the images reconstructed by different methods. It is obvious that the line intensity profile from our proposed method is closer to the ground truth, demonstrating that the proposed method allows the better reconstructing of initial energy density and target boundary.

We use three image quality indexes: root mean square error (RMSE), structural similarity index (SSIM),[66] and peak signal-to-noise ratio (PSNR) for quantitative analysis and the results are shown in Table 1. The proposed TAT-Net obtains lower RMSE and higher SSIM and PSNR than the DAS, which indicates that the TAT-Net has obvious advantages over the traditional method in reconstruction accuracy. Compared with the U-Net, the TAT-Net has a slight improvement in RMSE and PSNR, and a substantial improvement in SSIM, suggesting that the TAT-Net has better structural preservation. We also record the reconstruction times of different methods, and the average computing time of TAT-NET is about 8 times faster than that of DAS.

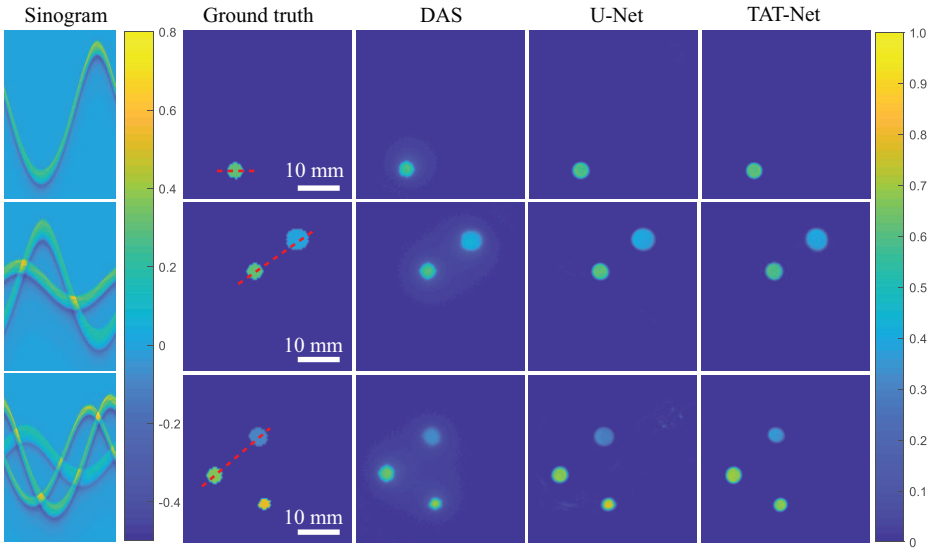


Fig. 3. Reconstructed images using synthetic data with 1 to 3 circular targets, where the first column denotes the the sinogram obtained from FEM simulation, the second column refers to the ground truth of energy density distribution, and the third to fifth column represent the images reconstructed using DAS, U-Net, and TAT-Net, respectively.

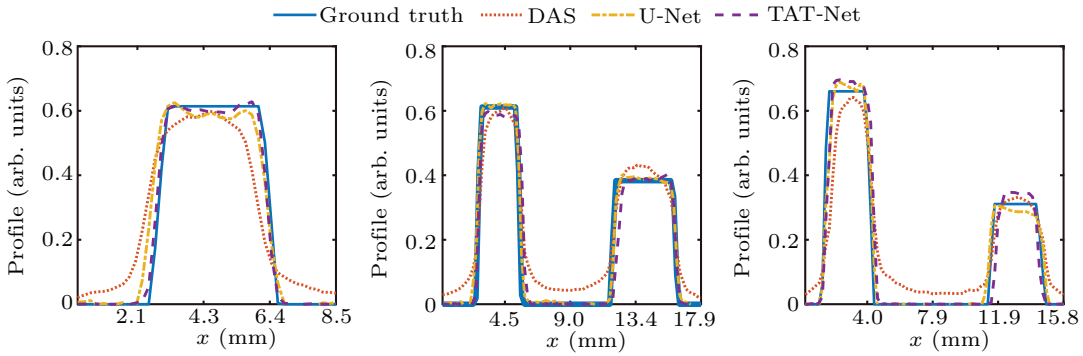


Fig. 4. Image profiles along the dashed red lines shown in the second column of Fig. 3.

Table 1. Comparison among different methods using simulation dataset.

Algorithms	RMSE	SSIM	PSNR	Time (s)
DAS	0.0242	0.723	33.93	0.244
U-Net	0.0142	0.960	38.61	0.025
TAT-NET	0.0143	0.988	38.64	0.031

3.2. Experimental data

In the experiments, one to three saline-containing tubes are used as reconstruction target(s), and the TAT phantom dataset is obtained by using the TAT system shown in Fig. 2. The network is trained on the synthetic dataset and tested on the TAT phantom dataset. Although both datasets contain similar circular reconstruction targets, significant domain gap can be observed in the sinogram due to the TAT system factors. These experiments are used to confirm whether the TAT-Net has the ability to generalize the experimental dataset.

Figure 5 shows the reconstruction results of 1 to 3 circular targets from different methods. The reconstructed images by the DAS suffer severe artifacts and blur target edges. There also exists an annular reflection signal interference around the target. In the case of 3 targets, the larger target is slightly elliptical and splits into two parts in the DAS image, and the shapes of two lager targets also become elliptical in the TAT-Net image. This is, we believe, due to inhomogeneous microwave absorption caused by microwave line polarization, resulting in a split elliptical shape of the reconstructed target.[67] Correspondingly, the sinogram of the larger target also splits into two lines. This type of data does not exist in our simulation training set, so the TAT-Net will experience distortion when reconstructing the data of such a type. A potential solution to this problem is to consider the signal change caused by microwave polarization in the training set simulation.

The U-Net can reconstruct the position and shape of each target, but severe artifacts appear, especially in the 3-target case. By observation, it can be obviously seen that these artifacts and sinogram have very similar characteristics, and we suspect that the appearance of sinogram-like artifacts are due to the shortcut connection structure in the U-Net, and the shallow connections in U-Net can give rise to the decrease of the generalization ability and robustness of the model. We note that the TAT-Net removes the shortcut link structure compared with U-Net and provides the artifact-free images, which demonstrates our hypothesis.

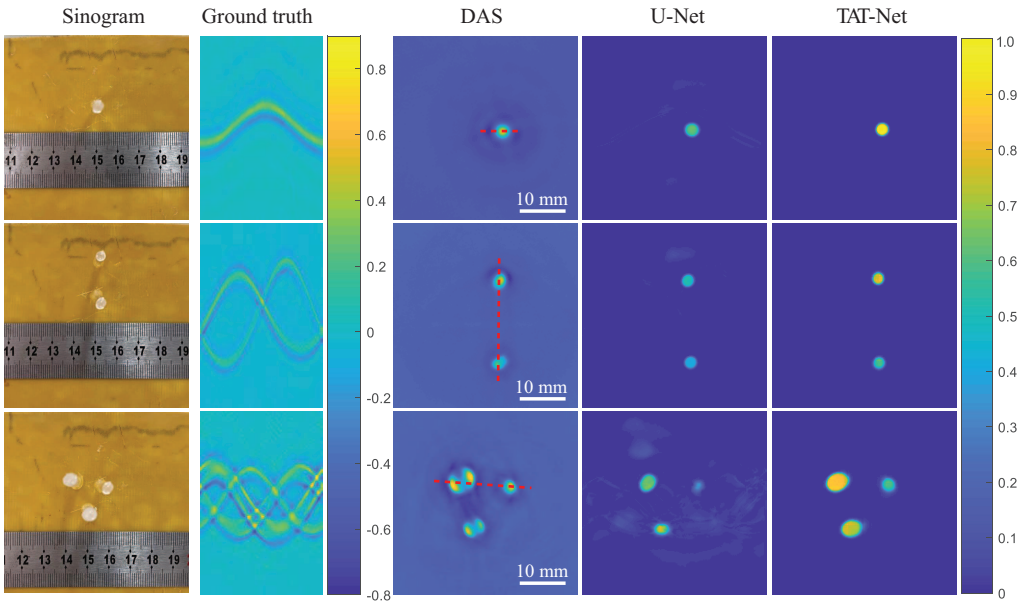


Fig. 5. Reconstructed image using experimental data with 1 to 3 circular targets, Where the first column denotes the photograph ofe phantoms, the second column represents sinograms obtained from experimental data, and the third to fifth column refer to the images reconstructed using DAS, U-Net, and TAT-Net, respectively.

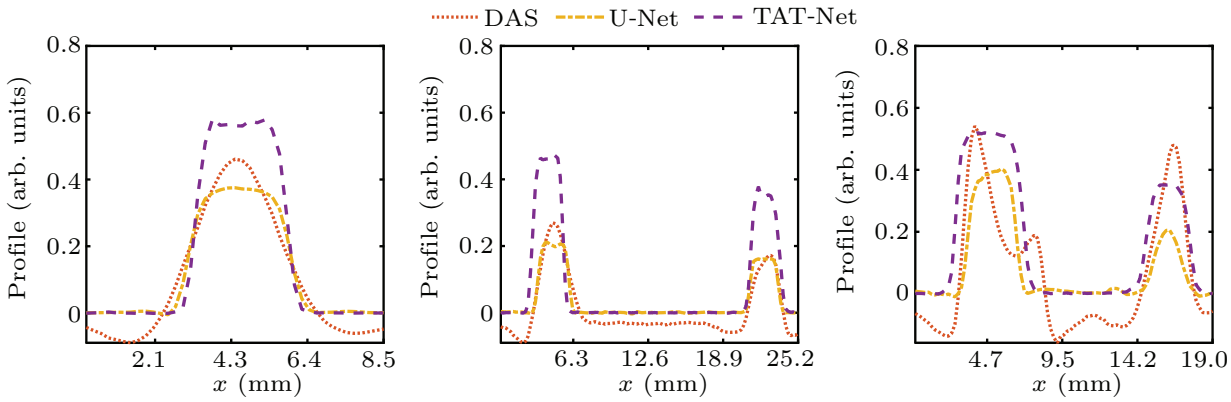


Fig. 6. Image profiles along dashed red lines shown in the second column of Fig. 5.

The 1D line intensity profiles shown in Fig. 6 clearly demonstrate the stability of TAT-Net against disturbances from TAT system. In addition, the TAT-Net obtains a correct target shape based on real characteristics, which shows great robustness against domain gap.

3.3. Robustness against number of targets

Since the training set only contains the data of 3 targets at most, we study the robustness of TAT-Net against more imaging targets. Figures 7 and 8 show the results on 4 and 5 circular targets in simulations and experiments, respectively.

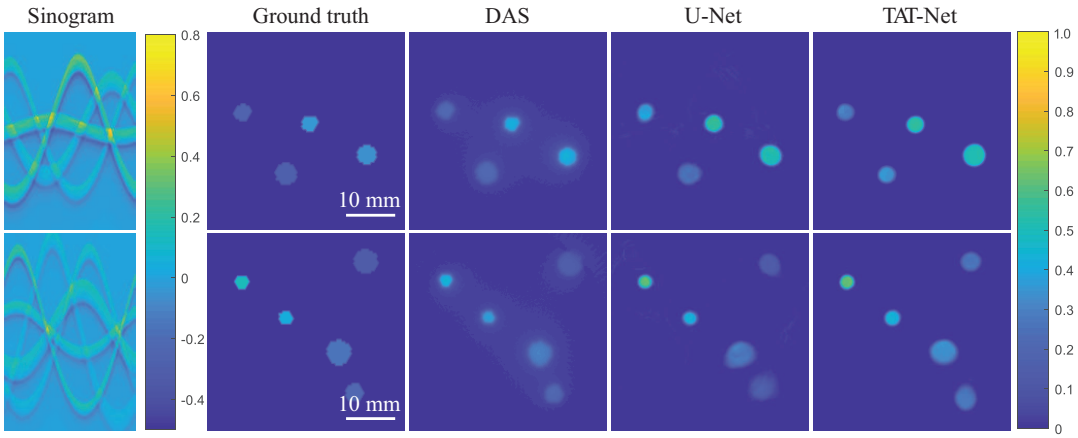


Fig. 7. Reconstructed images using simulation data with 4 and 5 circular targets.

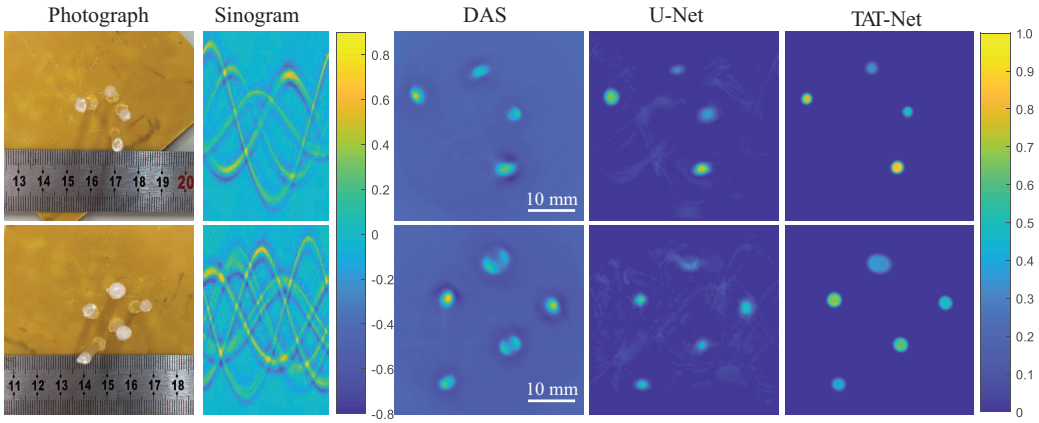


Fig. 8. Reconstructed images by using experimental data with 4 and 5 circular targets.

In simulation, the TAT-Net can completely reconstruct 4 and 5 circular targets, however, some circles have slight elliptical distortion and shift in the 5 circular case. Unlike the distortion in Fig. 5, this problem caused mainly by the generalization error of the model. In the TAT phantom experiments, it can be seen that the DAS is susceptible to artifacts that produce a blurred image and U-Net is almost impossible to reconstruct correctly. By contrast, the TAT-Net can still give a satisfactory quality of image reconstruction, demonstrating its great potential in the more complex situations.

3.4. Analysis of quantitative reconstruction

Most of the existing TAT reconstruction methods are qualitative, which means that the reconstructed image can only provide the structural characteristics and relative amplitude of the reconstruction target. The TAT-Net has the ability to be quantitatively reconstructed by taking the advantage of synthetic data, where the real initial energy density can be used as ground truth. In this subsection, we verify the ability to quantitatively reconstruct the energy loss density by using the TAT-Net. We study 64 sets of simulation samples and calculate the maximum value inside a reconstruction target as the initial energy loss density. Figure 9(a) shows the distributions of all the samples and the fitting line based on the least square method. It can be clearly seen that there is a very high correlation ($R = 0.9858$) between the TAT-Net and the ground truth; this demonstrates that the proposed method provides a stable and accurate quantitative energy density prediction.

In the experiment, one saline-containing tube is placed at different positions and we can obtain a series of phantom data with different energy density distributions due to the inhomogeneous microwave energy field. The real energy density cannot be measured, and thus, we take the reconstructed images by DAS as reference and calculate the correlation ($R = 0.9295$) between TAT-Net and DAS. Figure 9(b) gives the results of 90 sets of experimental samples, which indirectly demonstrate the ability of the proposed method to implement the quantitative reconstruction.

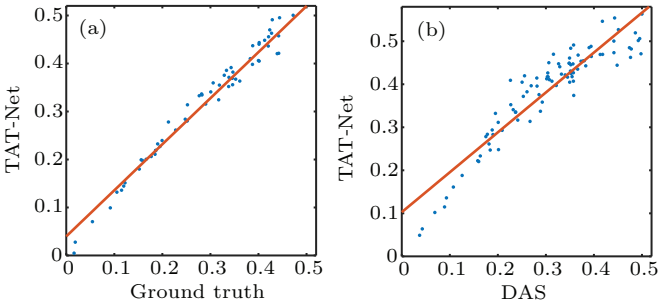


Fig. 9. (a) Plots of TAT-Net *versus* ground truth quantification, showing very high correlation between maximum values of reconstructed images and ground truth in one target ($R = 0.9858$); (b) plots of TAT-Net *versus* DAS quantifications, indicating also a high correlation between maximum values of reconstructed images with TAT-Net and DAS ($R = 0.9295$).

4. Discussion

In this work, we propose a novel direct-learning reconstruction method for thermoacoustic tomography, and design a CNN architecture called TAT-Net to reconstruct the initial energy density distribution from sinogram data. A more efficient up-sampling structure is used in the TAT-Net. This not only reduces the number of model parameters, but also improves expressive ability of the neural network. The finite element simulation method is used to generate a large quantity of simulation data to train the network, and the feasibility and stability of the network are validated by using the experimental data. Compared with traditional methods, the TAT-Net costs less time to obtain clearer images with fewer artifacts. Finally, we study the quantitative imaging capability of the proposed method, and the results show that the method can provide a reliable initial energy density distribution.

The network maps the sinogram domain to the image domain, and the network's capability set relies heavily on the training data. We need to consider two issues in future. (i) A new network will be trained if the parameters of the imaging system are changed, such as the type and position of the sensor, the parameters of microwave pulse source, the type of the imaging target, etc. One solution is to use image sam-

pling to match the resolution to the model input. The other approach is to add a large quantity of data to cover the range of variation of these parameters while requiring a more expressive and generalizable network. Our study suggests that our proposed method is suitable for an imaging system with stable parameters. Although this problem limits the applications in some experimental systems that require the parameters to be adjusted frequently, the proposed method has greater advantages for the systems or products with stable parameters. (ii) Networks trained with synthetic data may not be suitable for experimental data, therefore we propose a preprocessing scheme based on the signal processing and physical model for reducing the domain gap between synthetic data and experimental data. The experimental results demonstrate the reliability of the method. Our future research will focus on how to construct more applicable synthetic data in the case of *in-vivo* data.

5. Conclusions

We provide a new paradigm to reconstruct the TAT images with deep learning. In principle, the method is also applicable to other imaging methods such as the PAT. The TAT-Net takes the advantages of bottleneck residual structure and the efficient deconvolution up-sampling block. It also shows the ability to accurately reconstruct the quantitative initial energy density distribution in simulations and experiments. The finite element method demonstrates great potential for virtual data generation, which can be used in other imaging techniques such as the PAT. In the future we will investigate the performance of our method applied to *in-vivo* data.

References

[1] Zheng Z and Jiang H B 2019 *Quant. Imaging Med. Surg.* **9** 625
[2] Zheng Z, Huang L and Jiang H 2019 *International Applied Computational Electromagnetics Society Symposium*, August 8–11, 2019, Nanjing, China, 1–2
[3] Zheng Z, Jiang Y, Huang L, Zhao Y and Jiang H 2020 *J. Xray Sci. Technol.* **28** 137
[4] Chi Z, Zhao Y, Yang J, Li T, Zhang G and Jiang H 2019 *IEEE Trans. Biomed. Eng.* **66** 1598
[5] Huang L, Yao L, Liu L X, Rong J and Jiang H B 2012 *Appl. Phys. Lett.* **101** 244106
[6] Luo W, Ji Z, Yang S and Xing D 2018 *Phys. Rev. Appl.* **10** 024044
[7] Xu F, Ji Z, Chen Q, Yang S and Xing D 2019 *IEEE Trans. Med. Imaging* **38** 205
[8] Kruger R A, Kopecky K K, Aisen A M, Reinecke D R, Kruger G A and Kiser W L 1999 *Radiology* **211** 275
[9] Fu Y, Ji Z, Ding W, Ye F and Lou C 2014 *Med. Phys.* **41** 110701
[10] Ji Z, Lou C, Yang S and Xing D 2012 *Med. Phys.* **39** 6738
[11] Wang B, Zhao Z, Liu S, Nie Z and Liu Q 2017 *Appl. Phys. Lett.* **111** 223701
[12] Wang X, Bauer D R, Witte R and Xin H 2012 *IEEE Trans. Biomed. Eng.* **59** 2782
[13] Patch S K and See W A 2016 *Photons Plus Ultrasound: Imaging and Sensing*, March 15, 2016, San Francisco, USA
[14] Eckhart A T, Balmer R T, See W A and Patch S K 2011 *IEEE Trans. Biomed. Eng.* **58** 2238
[15] Huang Y, Kellnberger S, Sergiadis G and Ntziachristos V 2018 *Sci. Rep.* **8** 15522
[16] Zheng Z, Huang L and Jiang H 2018 *Appl. Phys. Lett.* **113** 253702
[17] Zhao Y, Chi Z, Huang L, Zheng Z, Yang J and Jiang H 2017 *Journal of Innovative Optical Health Sciences* **10** 1740001
[18] Mrozowski M, Okoniewski M, Okoniewska E and Stuchly M A 1997 *IEEE MTT-S International Microwave Symposium Digest*, June 8–13, 1997, Denver, USA, pp. 95–97
[19] Xu M and Wang L V 2005 *Phys. Rev. E* **71** 016706
[20] Haltmeier M 2013 *Comput. Math. Appl.* **65** 1025
[21] Burgholzer P, Bauer-Marschallinger J, Grün H, Haltmeier M and Paltauf G 2007 *Inverse Problems* **23** S65
[22] Cox B T and Treeby B E 2010 *IEEE Trans. Med. Imaging* **29** 387
[23] Treeby B E, Zhang E Z and Cox B T 2010 *Inverse Problems* **26** 115003
[24] Hristova Y, Kuchment P and Nguyen L 2008 *Inverse Problems* **24** 055006
[25] Yuan Z and Jiang H 2006 *Appl. Phys. Lett.* **88** 231101
[26] Belhachmi Z, Glatz T and Scherzer O 2016 *Inverse Problems* **32** 045005
[27] Huang C, Wang K, Nie L, Wang L V and Anastasio M A 2013 *IEEE Trans. Med. Imaging* **32** 1097
[28] Schwab J, Pereverzyev S and Haltmeier M 2018 *Siam Journal on Numerical Analysis* **56** 160
[29] Wang K, Su R, Oraevsky A A and Anastasio M A 2012 *Phys. Med. Biol.* **57** 5399
[30] Javaherian A and Holman S 2017 *IEEE Trans. Med. Imaging* **36** 696
[31] Arridge S R, Betcke M M, Cox B T, Lucka F and Treeby B E 2016 *Inverse Problems* **32** 115012
[32] Xu Y, Wang L V, Ambartsoumian G and Kuchment P 2004 *Med. Phys.* **31** 724
[33] Zhang Z, Liang X, Dong X, Xie Y and Cao G 2018 *IEEE Trans. Med. Imaging* **37** 1407
[34] Li Y, Li K, Zhang C, Montoya J and Chen G 2019 *IEEE Trans. Med. Imaging* **38** 2469
[35] Yin X, Zhao Q, Liu J, Yang W, Yang J, Quan G, Chen Y, Shu H, Luo L and Coatrieux J 2019 *IEEE Trans. Med. Imaging* **38** 2903
[36] Hammernik K, Würfl T, Pock T and Maier A 2017 *Bildverarbeitung für die Medizin*, March 1, 2017, Berlin, Germany, pp. 92–97
[37] Kang E, Min J and Ye J C 2017 *Med. Phys.* **44** 360
[38] Gupta H, Jin K H, Nguyen H Q, Mccann M T and Unser M 2018 *IEEE Trans. Med. Imaging* **37** 1440
[39] Han Y S, Yoo J and Ye J C 2016 arXiv: 1611.06391
[40] Jin K H, Mccann M T, Froustey E and Unser M 2017 *IEEE Trans. Image Process* **26** 4509
[41] Hammernik K, Klatzer T, Kobler E, Recht M P, Sodickson D K, Pock T and Knoll F 2018 *Magn. Reson. Med.* **79** 3055
[42] Han Y, Yoo J, Kim H H, Shin H J, Sung K and Ye J C 2018 *Magn. Reson. Med.* **80** 1189
[43] Xiang L, Chen Y, Chang W, Zhan Y, Lin W, Wang Q and Shen D 2018 *IEEE Trans. Biomed. Eng.* **66** 2105
[44] Allman D, Reiter A and Bell M a L 2018 *IEEE Trans. Med. Imaging* **37** 1464
[45] Antholzer S, Haltmeier M, Nuster R and Schwab J 2018 *Photons Plus Ultrasound: Imaging and Sensing*, 2018, San Francisco, USA, 104944U
[46] Boink Y E, Manohar S and Brune C 2019 arXiv: 1906.07499
[47] Johnstonbaugh K, Agrawal S, Abhishek D, Homewood M, Karri S P K and Kothapalli S R 2019 *Photons Plus Ultrasound: Imaging and Sensing*, 2019, San Francisco, USA, 108781L
[48] Allman D, Reiter A and Bell M a L 2017 *IEEE International Ultrasonics Symposium*, September 6–9, 2017, Washington, USA, pp. 1–4
[49] Lan H, Yang C, Jiang D and Gao F 2019 *IEEE International Ultrasonics Symposium*, October 6–9, 2019, Glasgow, Scotland, pp. 487–489
[50] Kim K, Wu D, Gong K, Dutta J, Kim J H, Son Y D, Kim H K, Fakhri G E and Li Q 2018 *IEEE Trans. Med. Imaging* **37** 1478
[51] Waibel D, Groehl J, Isensee F, Kirchner T, Maier-Hein K and Maier-Hein L 2018 *Photons Plus Ultrasound: Imaging and Sensing*, 2018, San Francisco, USA, 104942S
[52] Anas E M A, Zhang H K, Audigier C and Boctor E M 2018 *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, September 16–20, 2018, Granada, Spain, pp. 3–11
[53] Lan H, Yang C, Jiang D and Gao F 2019 41*st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, July 23–27, 2019, Berlin, Germany, pp. 7115–7118

[54] Lan H, Zhou K, Yang C, Cheng J, Liu J, Gao S and Gao F 2019 *Medical Image Computing and Computer Assisted Intervention*, October 13–17, 2019, Shenzhen, China, pp. 273–281

[55] Lan H, Zhou K, Yang C, Liu J, Gao S and Gao F 2019 41*st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, July 23–27, 2019, Berlin, Germany, pp. 6367–6370

[56] Hauptmann A, Lucka F, Betcke M, Huynh N, Adler J, Cox B, Beard P, Ourselin S and Arridge S 2018 *IEEE Trans. Med. Imaging* **37** 1382

[57] Aggarwal H K, Mani M P and Jacob M 2019 *IEEE Trans. Med. Imaging* **38** 394

[58] Odena A, Dumoulin V and Olah C 2016 *Distill* **1**

[59] Wang X, Yu K, Wu S, Gu J, Liu Y, Dong C, Qiao Y and Loy C C 2019 *Computer Vision - ECCV* 2018 *Workshops*, September, 8–14, 2019, Munich, Germany, pp. 63–79

[60] He K, Zhang X, Ren S and Sun J 2016 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 27–30, 2016, Las Vegas, USA, pp. 770–778

[61] He K, Zhang X, Ren S and Sun J 2015 *IEEE International Conference on Computer Vision*, December 7–13, 2015, Santiago, Chile, pp. 1026–1034

[62] Kingma D P and Ba J 2014 arXiv: 1412.6980

[63] Paszke A, Gross S, Massa F, *et al.* 2019 *Adv. Neural Inf. Process. Syst.* **32** 8026

[64] Yao L, Guo G and Jiang H 2010 *Med. Phys.* **37** 3752

[65] Wang B, Xiong W, Su T, Xiao J and Peng K 2018 *Appl. Opt.* **57** 9123

[66] Zhou W, Bovik A C, Sheikh H R and Simoncelli E P 2004 *IEEE Trans. Image Process.* **13** 600

[67] He Y, Shen Y C, Liu C J and Wang L H V 2017 *Appl. Phys. Lett.* **110** 053701