Deep Neural Network based Bandwidth Enhancement of Photoacoustic Data

Sreedevi Gutta[a], **Venkata Suryanarayana Kadimesetty**[a], **Sandeep Kumar Kalva**[b], **Manojit Pramanik**[b], **Sriram Ganapathy**[c], **Phaneendra K. Yalavarthy**[a,*]
[a]Department of Computational and Data Sciences, Indian Institute of Science, Bangalore-560 012 India
[b]Nanyang Technological University, School of Chemical and Biomedical Engineering, 62 Nanyang Drive, Singapore
[c]Department of Electrical Engineering, Indian Institute of Science, Bangalore-560 012 India

Abstract. Photoacoustic (PA) signals collected at the boundary of tissue are always band-limited. A deep neural network (DNN) was proposed to enhance the bandwidth of the detected PA signal, thereby improving the quantitative accuracy of the reconstructed PA images. A least square based deconvolution method that utilizes the Tikhonov regularization framework was used for comparison with the proposed network. The proposed method was evaluated using both numerical and experimental data. The results indicate that the proposed method was capable of enhancing the bandwidth of the detected PA signal, which inturn improves the contrast recovery and quality of reconstructed PA images without adding any significant computational burden.

Keywords: photoacoustic data; deep neural network; bandwidth enhancement; reconstruction.

*Address all correspondence to: Phaneendra K. Yalavarthy, E-mail: yalavarthy@iisc.ac.in

Photoacoustic tomography (PAT) is a noninvasive and hybrid biomedical imaging modality that combines the merits of both optical and ultrasound imaging techniques.[1–5] In this modality, a nanosecond duration short laser pulse is used to illuminate the tissue of interest. The deposited energy is absorbed by the tissue resulting in a small temperature rise in the order of milliKelvin. Subsequently, acoustic waves are generated due to thermoelastic expansion and are collected using ultrasonic transducers placed on the boundary of the tissue. The detected photoacoustic (PA) signal is used to map the initial pressure with the help of reconstruction algorithms. The initial pressure is proportional to the absorption coefficient of the tissue, assuming homogeneous fluence distribution.

The reconstruction algorithms estimate the initial pressure distribution from the measured PA signal using either analytical or model based iterative techniques.[6] The performance of the reconstruction algorithms depends on the quality of PA signal. The factors that effect this performance other than signal to noise ratio are limited bandwidth (BW) of acoustic detectors.[7,8] The limited bandwidth of the transducer results in the spatially invariant blurring of the reconstructed PA images. The signal received at the acoustic detectors is always band-limited. As it is unavoidable to use band-limited transducers, an additional step to restore the original PA signal will improve the quality of the reconstructed PA images. To have real-time imaging, it is also necessary that this additional step is computationally efficient.

The PA signal detected using the band-limited transducer can be simulated by convolving the original PA signal with the transducer impulse response.[9] Previously, deconvolution algorithms were proposed to overcome the bandwidth limitation of the ultrasound transducer.[9] These algorithms attempt to restore the original PA signal by deconvolving the detected PA signal with the transducer impulse response, which can be estimated either experimentally or can be obtained from the manufacturer. The deconvolution algorithms are in general ill-posed, necessitating the need of regularization to obtain a unique solution.[10] Recently, a method to recover out-of-band frequencies was proposed for PA imaging. This method was based on inverse (Wiener) filtering and it requires a best estimate of the frequency response of the ultrasound transducers.[11]

Application of deep convolutional neural network (DCNN) for natural image deconvolution was proposed recently by Xu. et. al [12], earlier to this it was applied to astronomical image reconstruction,[13] semantic segmentation,[14] and music classification.[15]

In this letter, a simple deep neural network (DNN) was proposed to enhance the bandwidth of the detected PA signal, which reconstructs all frequency components. The network was trained with band-limited signal as input and output being full bandwidth signal. The enhanced/predicted acoustic (PA) signals were used as input to the analytical reconstruction algorithms like back-projection to obtain quantitatively accurate PA images. Similar framework could be applicable to ultrasound imaging (including ultrasound tomography) to enhance the bandwidth of the recorded echo signal.

The photoacoustic wave propagation is modeled using the following differential equation:[6]

$$\nabla^2 P(x,t) - \frac{1}{v_s^2}\frac{\partial^2 P(x,t)}{\partial t^2} = -\Gamma \frac{\partial H(x,t)}{\partial t}\,, \tag{1}$$

where v_s represents the speed of sound in the medium, Γ denotes the Gruneisen coefficient, and $H(x,t)$ represents the absorbed energy per unit time per unit volume with x indicating the spatial location. The above equation (Eq. 1) can be equivalently written as:

$$\nabla^2 P(x,t) - \frac{1}{v_s^2}\frac{\partial^2 P(x,t)}{\partial t^2} = 0, \tag{2}$$

with initial conditions $\partial P/\partial t|_{t=0} = 0$ and $P|_{t=0} = \Gamma H(x)$. The acoustic wave equation can be discretized as a linear system of equations reducing to a matrix form as[16,17]

$$\mathbf{A}P_0 = b, \tag{3}$$

with the system matrix $\mathbf{A}$ containing the impulse responses of all the pixels in the imaging grid, P_0 indicates the initial pressure distribution, and b is the band-limited PA data collected by the transducers. Typically, Green's function approach is deployed to measure the impulse response for its computational efficiency.[18] The back-projection (analytical) type image reconstruction scheme becomes[16,17]

$$P_{bp} = \mathbf{A}^T b, \tag{4}$$

where $\mathbf{A}^T$ indicate the transpose of the matrix. The acquired PA signal in the presence of noise can be written as

$$b = \bar{b} * h + noise, \tag{5}$$

where $\bar{b}$ denotes the original PA signal and h indicates the impulse response of the transducer. A least square deconvolution with Tikhonov regularization was proposed earlier to remove the effect due to finite bandwidth of the transducer.[9] Equation 5 can be rewritten as

$$b = H\bar{b} + n, \tag{6}$$

where H is the circulant convolution matrix formed from h. The original PA signal can be obtained by minimizing the cost function given as

$$C(\bar{b}) = ||b - H\bar{b}||^2 + \lambda||\bar{b}||^2, \tag{7}$$

with λ being the regularization parameter. The minimum for the cost function is obtained by equating the gradient of Eq. 7 to zero,

$$H^T b = H^T H\bar{b}_{Tik} + \lambda\bar{b}_{Tik}, \tag{8}$$

where T indicates the transpose of a matrix and $\bar{b}_{Tik}$ is a least square deconvoluted signal. Equation 8 can be written in time domain as

$$h[-n] * b[n] = h[-n] * h[n] * \bar{b}_{Tik}[n] + \lambda\bar{b}_{Tik}[n], \tag{9}$$

where n denotes the time index. In frequency domain, Eq. 9 becomes

$$conj(H)B = conj(H)H\bar{B}_{Tik} + \lambda\bar{B}_{Tik}, \tag{10}$$

where B represents the Fourier transformed version of b. $\bar{B}_{Tik}$ from Eq. 10 is evaluated as

$$\bar{B}_{Tik} = \frac{conj(H)B}{conj(H)H + \lambda}. \tag{11}$$

The time domain signal $\bar{b}_{Tik}$, can be obtained by applying inverse Fourier transform to $\bar{B}_{Tik}$. The corrected PA signal $\bar{b}_{Tik}$ is then used for image reconstruction using Eq. 4.

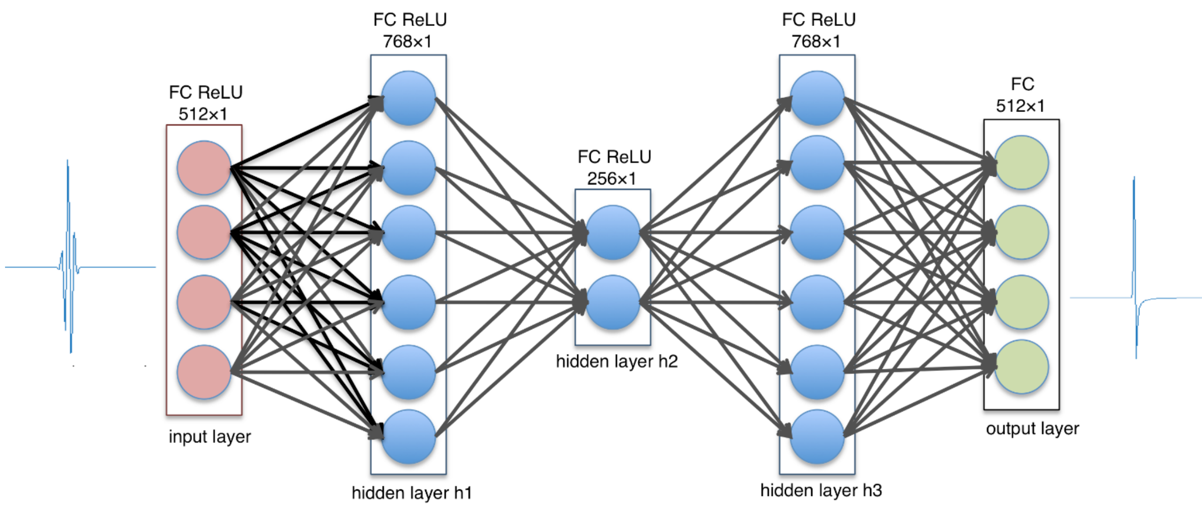


Fig 1 The neural network architecture listing the layers utilized in the proposed deep neural network. The network takes band-limited signal as input and predicts full bandwidth signal. FC – Fully Connected. ReLU – Rectified Linear Unit .

Recently, deep learning has been proven as a promising technique for many visual recognition tasks.[19,20] Deep learning based improvements to medical imaging has been discussed widely and known to have huge potential especially in enhancing the raw data.[21] The deep feedforward neural networks, extract the features through the cascade of several nonlinear activation functions (Fig.

1). As the enhancement of bandwidth of the collected boundary data can be seen as extrapolation of a nonlinear function, the approach followed in this work is to utilize the deep neural network in predicting the missing frequencies. This inturn removes the effect of finite bandwidth of transducers. The idea of the proposed network is to formulate the deconvolution process as a nonlinear regression.

The input to the network is a band-limited signal $x = (x_1, x_2, ..., x_N)$ and the output is a full bandwidth signal $y = (y_1, y_2, ..., y_N)$ where $x_i \& y_i \in \mathbb{R}$. Given a training dataset $(x_i, y_i)_{i=1}^{N}$, our goal is to determine a non-linear function ϕ that can effectively map between the input band-limited signal x and the output full bandwidth signal y: $y = \phi(x)$. The network architecture was shown in Fig. 1. The proposed network contains 5 fully connected (FC) layers, one being the input layer other being the output layer and the rest three are hidden layers. Our network architecture is similar to the decoder network.[22] Note that the encoder decoder network was proven to be efficient in the deep learning literature and has been applied in many scenarios like predicting the missing data.[22] The first and third hidden layers (h_1 and h_3 in Fig. 1) is designed with 768 (512+$\frac{512}{2}$) nodes, i.e, by increasing the original number of nodes by 50 %. The second hidden layer (h_2 in Fig. 1) contains 256 ($\frac{512}{2}$) nodes, i.e., considering half of the original nodes. The k-th layer of the network evaluates: $h_k = f(W_k h_{k-1} + b_k)$, where h_k indicates the k-th hidden layer, f is the activation function, W_k represents the matrix of trainable weights, and b_k is a vector of trainable biases. The activation function utilized in this network except for the last layer (as it is a regression problem) is rectified linear unit (ReLU). It should be noted that ReLU's are computationally efficient as they do not require to compute any exponentiation and division.[23]

The network was trained with mean squared error (MSE) as a loss function. The MSE can be written as

$$MSE = \frac{1}{N}\sum_{i=1}^{N} ||\hat{y}_i - \phi(x_i)||^2, \tag{12}$$

where $\hat{y}_i$ represents the predictions of the model on the training set. The optimal weights for the network are determined in such a way that the MSE reduces by observing a training set (x_i, y_i). In other words, the goal of training is to learn the model parameters weights and biases by minimizing MSE (cost function), which measures the discrepancy between the target (expected) and the predicted output. An efficient ADAM optimization algorithm[24] that optimizes stochastic objective functions was used to find the optimal weights. It combines the merits of AdaGrad[25] which is efficient for sparse gradients and RMSProp[26] known to be efficient for non-stationary objectives. Note that this method only requires evaluation of first-order gradients.

A computational grid of dimension 100 mm $\times$ 100 mm containing 1000 $\times$ 1000 pixels was considered. The data was generated for the objects of size 400 $\times$ 400. Hundred point detectors of center frequency 2.25 MHz and 70% bandwidth were placed around the tissue surface equidistantly on a circle of radius 37.02 mm. The data was recorded for a sampling frequency of 12.5 MHz with a total of 512 time steps. The medium was assumed to be homogeneous with uniform speed of sound as 1500 m/s and with no absorption and dispersion.

The proposed model is also validated using the experimental data. The schematic representation of the PAT imaging system used for performing the experiment is shown in Fig. 1(e) of Ref. 27. A Q-switched Nd:YAG laser (Continuum, Surelite Ex) delivering laser pulses of 532 nm wavelength of 5 ns duration at 10 Hz repetition rate was used. Four right-angle uncoated prisms (PS911, Thorlabs) and one uncoated planoconcave lens (LC1715, Thorlabs) were used to incident laser energy of density ~ 9 mJ/cm^2 (< 20 mJ/cm^2 : ANSI safety limit[28]) on the phantom. A triangular shaped horse hair phantom (side-length of hair: 10 mm and diameter of hair: 0.15 mm) data was collected at 100 locations around the sample using an ultrasound transducer (Olympus NDT, V306-SU) of 2.25 MHz center frequency with 13 mm diameter active area and 70% bandwidth. Another experimental phantom, circular in shape, made using low density polyethylene (LDPE) tubes (5 mm inner diameter) filled with black Indian ink was also utilized to evaluate the proposed method. The tubes were placed at 0 and 15 mm from the scanning center and affixed at the bottom of the acrylic slab. The detected PA signals were first amplified and filtered using a pulse amplifier (Olympus-NDT, 5072PR) and then recorded using a data acquisition (DAQ) card (GaGe, compuscope 4227) inside a desktop (Intel Xeon 3.7 GHz 64-bit processor, 16 GB RAM, running windows 10 operating system). A TTL sync signal from the laser was used to synchronize the data acquisition with laser illumination. The original data was collected at a sampling frequency of 25 MHz containing 1024 time points and simulations were performed at a sampling frequency of 12.5 MHz as the original data was subsampled to 512 time points. Figures of merit, such as Pearson's correlation coefficient (PC), contrast to noise ratio (CNR), and signal to noise ratio (SNR) were used to quantitatively show the improvement obtained from the proposed method and are defined in Refs. [17, 27].

For training of the network, numerical breast phantoms[29] were used to obtain 286300 (generated using 2863 phantoms, obtained via considering two-dimensional slices in the Z-direction of these breast phantoms, with 100 detector positions for each slice) samples of full bandwidth and limited bandwidth PA signals. The breast phantoms were chosen for training as they provide close realization of patient anatomical structures, thus enabling realistic numerical experiments. An open source k-wave toolbox was utilized to generate the data from these phantoms.[30] The network was trained for a total of 100 epochs on a batch size of 100. The validation was performed

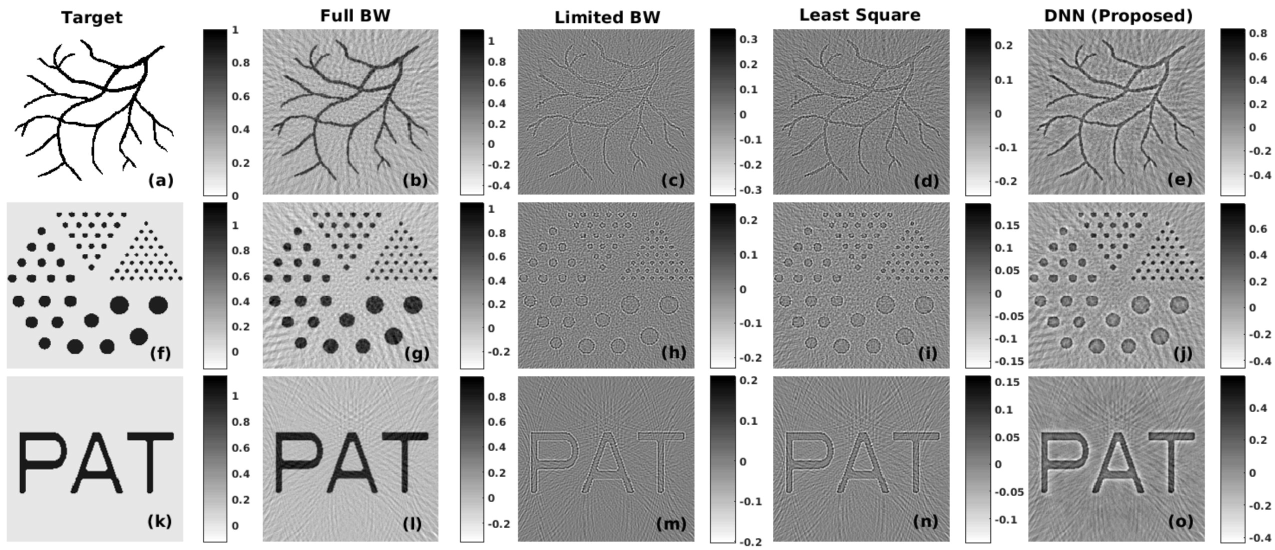


Fig 2 Numerical phantoms used in evaluating this work: (a) Blood vessel network, (f) Derenzo phantom, and (k) PAT phantom. Reconstructed (back-projected) initial pressure images with 100 detectors using (b,g,l) full BW signal, (c,h,m) limited BW signal, (d,i,n) predicted signal from least square deconvolution method, and (e,j,o) predicted signal from the proposed deep neural network (DNN). The SNR of the data is at 40 dB.

on 1% of the training data. Keras software (https://keras.io/), with backend as Theano (http://deeplearning.net/software/theano/), which is an open source python library was utilized for training of the proposed network. The training loss in the end of 100 epochs is 2.0826e-4. All computations were performed on a Linux workstation with dual six-core Intel Xeon processor having a speed of 2.66 GHz with 64 GB RAM. To evaluate the degree of correlation between the target and the reconstructed image, PC was computed. It is given as[17]

$$PC(P_0, P_{recon}) = \frac{cov(P_0, P_{recon})}{\sigma(P_0)\sigma(P_{recon})}, \tag{13}$$

where P_0 represents the target initial pressure distribution, P_{recon} is the reconstructed initial pressure distribution, σ is the standard deviation, and cov is the covariance.

To measure the differentiability of the region of interest with respect to background, CNR was utilized. It can be defined as[17]

$$CNR = \frac{\mu_{roi} - \mu_{back}}{(\sigma_{roi}^2 a_{roi} + \sigma_{back}^2 a_{back})^{1/2}}, \tag{14}$$

where μ and σ corresponds to mean and standard deviation respectively. The $a_{roi} = \frac{A_{roi}}{A_{total}}$, where A_{roi} indicates the number of pixels with non zero initial pressure rise in the target and A_{total} is for the total number of pixels in the imaging grid, $a_{back} = \frac{A_{back}}{A_{total}}$ with A_{back} denoting the number of pixels with zero initial pressure rise in the target.

For the experimental case, where it is not possible to determine the target initial pressure distribution, SNR was computed. It is expressed as

$$SNR(dB) = 20 \times log_{10}\left(\frac{S}{n}\right), \tag{15}$$

with S denoting the peak-to-peak signal amplitude and n corresponds to standard deviation of noise. Three numerical phantoms (different from the training data) as shown in the first column of Fig. 2 were considered to evaluate the effectiveness of the proposed method: (a) Blood vessel network is generally used as PA numerical phantom for imaging blood vasculature, (b) Derenzo phantom containing various size circular pressure distribution, and (c) PAT phantom to simulate sharp edges. The PA time domain and frequency response plots for full BW, limited BW, predicted data from least square deconvolution technique, and predicted data from DNN were shown in Fig. 3. Figure 3(a-b) corresponds to the time domain and frequency spectrum plot for the blood vessel phantom shown in Fig. 2(a) using the transducer located at 9 o'clock position. The time domain and amplitude spectrum plots for the derenzo Fig. 2(f) and PAT phantom Fig. 2(k) for

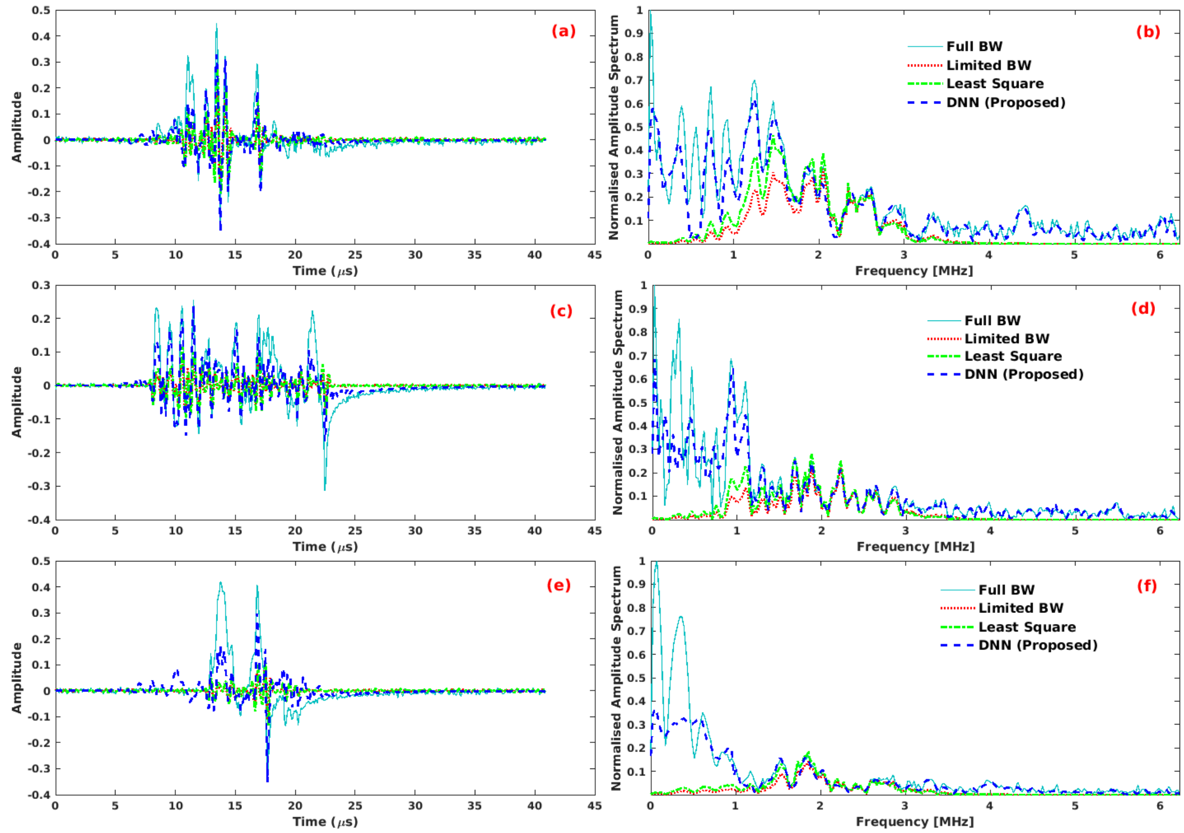


Fig 3 The PA signal time domain and frequency response plot for the data collected by a single transducer. (a-b) Blood vessel phantom shown in Fig. 2(a), (c-d) Derenzo phantom shown in Fig. 2(f), and (e-f) PAT phantom shown in Fig. 2(k).

the data collected using the same transducer were given in Fig. 3(c-d) and (e-f) respectively. The bandwidth enhancement using the proposed DNN can be observed from these plots (especially from the frequency response plots), where frequency response of the signal estimated using the proposed method was more close to full BW signal response.

Due to the limited BW of the transducers, the signal collected by the transducers is highly attenuated below 0.6 MHz and above 3.5 MHz. It can be noted that the removal of low frequencies has significant effect on the signal amplitude. It was evident that the proposed DNN was better able to predict the missing frequencies than the least square deconvolution method. In this case, originally collected photoacoustic signals are band-pass filtered, the high frequency components are generally more attenuated.[11] As the proposed network is formulated as a deconvolution process, which essentially recovers the high frequency components (equivalent of removing the blur), it can be noticed from Fig. 3 that the proposed DNN method was recovering higher frequency components close to expected signals compared to lower frequency components.

The reconstructed images using the full BW signal were shown in the second column of Fig. 2. The third column of Fig. 2 corresponds to the reconstructed images for limited BW signal. The reconstruction results for the data predicted using least square deconvolution method with $\lambda = 0.1$ (heuristic choice as reported in Ref. 9), were presented in fourth column of Fig. 2. The last column of Fig. 2 corresponds to the data predicted using the proposed deep neural network. The quantitative metrics for these results were given in Table 1. It was clearly evident from Fig. 2 that the analytical back-projection using band-limited signal was insufficient in recovering the original object. Previously, the model based algorithms as proposed in Ref. 17 were shown to recover the original object shape, but these methods were computationally expensive. It can be seen that the analytical reconstruction using enhanced bandwidth PA signal from DNN was able to provide atleast three times better quantitative accuracy compared to the state of the art deconvolution method. The quantitative accuracy can be determined from the initial pressure rise values as shown in target phantoms. Note that the proposed network output in the training is noiseless full BW signal, the predicted signals using the trained network will be a closer approximation to the expected noiseless signal. Thus denoising step is inherently embedded in the proposed method.

As the network was trained with blood vessel type of phantoms, the quantitative accuracy obtained from the proposed method was superior for blood vessel phantom compared to the other numerical phantoms (derenzo and PAT). Increasing the training set with wide range of phantoms can further improve the reconstruction results. Training of the proposed network took around 3

Table 1 The quantitative metrics, PC and CNR of the reconstructed images presented in Fig. 2.

Numerical Experiments	Blood Vessel		Derenzo		PAT	
	PC	CNR	PC	CNR	PC	CNR
Full BW	0.85	5.74	0.92	5.03	0.96	9.89
Limited BW	0.30	1.14	0.26	0.05	0.22	0.01
Least square deconvolution	0.38	1.44	0.36	0.17	0.32	0.29
DNN (Proposed)	0.78	4.22	0.70	1.92	0.75	2.54

hours and it should be noted that the training was done only once for specific detection parameters. The training time could be further decreased (as low as 15 minutes) with utilization of general purpose graphics processing units (GPUs). The computational time for reconstructing the initial pressure distribution using back projection was 1.05 seconds. In addition, predicting the data using least square deconvolution method requires 80.58 milli seconds, whereas the proposed DNN takes 49.48 milli seconds. The proposed network was almost 1.63 times faster than the state of the art least square deconvolution method.

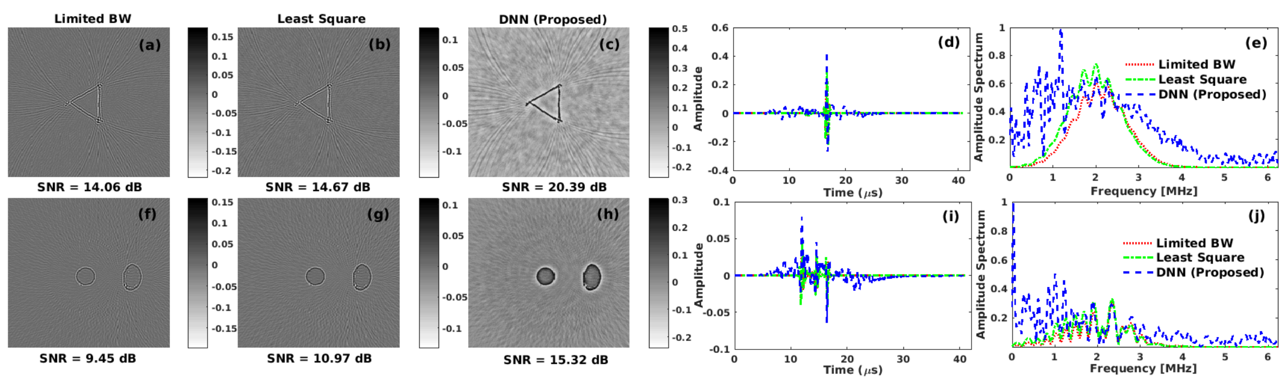


Fig 4 Reconstructed (back-projected) initial pressure images for horse hair phantom and ink tube phantom using (a,f) limited BW signal, (b,g) predicted signal from least square deconvolution method, and (c,h) predicted signal from the proposed deep neural network (DNN). Corresponding estimated SNR is specified at the bottom of each image. The photoacoustic signal time domain and frequency response plot for the transducer located at 9 o'clock position for (d,e) horse hair phantom (i,j) ink tube phantom.

The experimental results for the horse hair phantom data were shown in the first row of Fig. 4. The reconstructed image using the limited BW data was shown in Fig. 4(a). Figures 4(b) and (c) correspond to the reconstructed images for the data predicted using the least square deconvolution method with $\lambda = 0.1$ and the proposed DNN respectively. Representative time domain signal and frequency response plots for the experimental horse hair phantom data were given in Fig. 4(d-e) corresponding to the results shown in Fig. 4(a-c). The results for the ink tube phantom were given in the second row of Fig. 4. The initial pressure distribution for the limited BW, predicted data from least square deconvolution, and the proposed DNN methods were shown in Fig. 4(f-h) respectively. The corresponding representative time domain signal and frequency response plots were shown in Fig. 4(i-j). The proposed DNN provides enhanced time domain response compared to its counterpart. As the initial pressure rise was unavailable for the experimental data, the SNR was computed and specified below each image. It can be seen that the proposed method was able to provide better SNR images compared to the state of the art deconvolution method. The least square deconvolution method works well in the complete data case (close to 200 detectors for experiments similar to the ones conducted here[9]), whereas only limited data case (with only 100 detectors) was considered in this work, due to which the performance is comparatively limited.

The performance of the deep learning model is typically constrained by the quality of the training data. Deep learning techniques perform better with large amounts of data. It was shown in the literature[31] that maintaining diversity among the training samples can have significant effect on the outcome of the training. For the proposed method to provide best performance, the training data distribution should include the numerical and experimental cases presented here. But, in here, even though this was not fulfilled for all cases (especially Derenzo, PAT, and experimental ink tube phantom), the model was still able to provide superior performance in terms of enhancing the bandwidth of the PA signal (as evidenced by the figures of merit reported in Table-1 and Fig. 4).

The utilization of the band-limited signals without any pre-processing in the image reconstruction procedure results in loss of quantitative accuracy. As experimentally available transducers have only finite bandwidth, the signals that are recorded will always be band-limited. The proposed DNN helps to predict the full bandwidth signal from the band-limited signal and is proven to be effective compared to the state of the art methods. It was also shown through the numerical and experimental phantom experiments that using the predicted signals from DNN in the reconstruction procedure (back projection) results in better contrast recovery as well as improved PA image quality. In summary, as commercially available ultrasound detectors are bandwidth limited, typically having only 70% bandwidth, the proposed method can be seen as a promising pre-processing method to perform the bandwidth enhancement without adding any significant computational burden. These methods are critical for making the photoacoustic imaging more appealing for real-time imaging studies. The developed python code was provided as an open-source to the enthusiastic users.[32]

Disclosures

The authors have no financial interests or conflicts of interest in this manuscript.

Acknowledgments

This work was supported by Department of Biotechnology (DBT) Innovative Young Biotechnologist Award (IYBA) (Grant No: BT/07/IYBA/2013-13) and DBT Bioengineering Grant (No: BT/PR7994/MED/32/284/2013). The authors would also like to acknowledge financial support from the Tier 1 grant funded by the Ministry of Education in Singapore (RG41/14: M4011285, RG48/16: M4011617), and the Singapore National Research Foundation administered by the Singapore Ministry of Healths National Medical Research Council (NMRC/OFIRG/0005/2016: M4062012).

References

1 L. H. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).

2 L. Li et al., "Single-impulse panoramic photoacoustic computed tomography of small- animal whole-body dynamics at high spatiotemporal resolution," *Nature Biomedical Engineering* **1**(5), 0071 (2017).

3 P. K. Upputuri and M. Pramanik, "Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review," *J. Biomed. Opt.* **22**(4), 041006 (2017).

4 L. V. Wang and J. Yao, "A practical guide to photoacoustic tomography in the life sciences," *Nat. Methods* **13**(8), 627 (2016).

5 Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *J. Biomed. Opt.* **21**(6), 061007 (2016).

6 A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in optoacoustic tomography: A review," *Curr. Med. Imaging Rev.* **9**(4), 318–336 (2013).

7 M. Haltmeier and G. Zangerl, "Spatial resolution in photoacoustic tomography: effects of detector size and detector bandwidth," *Inverse Probloms* **26**(12), 125002 (2010).

8 M. Xu and L. H. V. Wang, "Analytic explanation of spatial resolution related to bandwidth and detector aperture size in thermoacoustic or photoacoustic reconstruction," *Phys. Rev. E* **67**(5), 056605 (2003).

9 N. A. Rejesh, H. Pullagurla, and M. Pramanik, "Deconvolution based deblurring of reconstructed images in photoacoustic/thermoacoustic tomography," *J. Opt. Soc. Am. A* **30**(10), 1994–2001 (2013).

10 S. M. Riad, "The deconvolution problem: an overview," *in Proceedings of the IEEE* **74**(1), 82–85 (1986).

11 M. Cao et al., "Spread Spectrum Photoacoustic Tomography With Image Optimization," *IEEE Trans. Biomed. Circuits Syst.* **11**(2), 411–419 (2017).

12 L. Xu et al., "Deep convolutional neural network for image deconvolution," *Advances in Neural Information Processing Systems*, 1790–1798 (2017).

13 R. Flamary, "Astronomical image reconstruction with convolutional neural networks," arXiv:1612.04526 (2016).

14 H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," *In Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 1520–1528 (2015).

15 K. Choi, G. Fazekas, and M. Sandler, "Explaining deep convolutional neural networks on music classification," arXiv:1607.02444 (2016).

16 C. B. Shaw et al., "Least squares QR-based decomposition provides an efficient way of computing optimal regularization parameter in photoacoustic tomography," *J. Biomed. Opt.* **18**(8), 80501 (2013).

17 J. Prakash et al., "Basis pursuit deconvolution for improving model-based reconstructed images in photoacoustic tomography," *Biomed. Opt. Express* **5**(5), 1363–1377 (2014).

18 E. Galetti, D. Halliday, and A. Curtis, "A simple and exact acoustic wavefield modeling code for data processing, imaging, and interferometry applications," *Geophysics* **78**, F17 (2013).

19 R. Girshick et al., "Rich feature hierarchies for accurate object detection and semantic segmen-

tation," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 580–587 (2014).

20 A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *In Advances in neural information processing systems*, 1097–1105 (2012).

21 G. Wang, "A perspective on deep imaging," *IEEE Access* **4**, 8914–8924 (2016).

22 V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," arXiv preprint arXiv:1511.00561 (2015).

23 M. D. Zeiler et al., "On rectified linear units for speech processing," *In Acoustics, Speech and Signal Processing (ICASSP)*, 3517–3521 (2013).

24 D. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).

25 J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research* **12**, 2121–2159 (2011).

26 T. Tieleman T and G. Hinton, Lecture 6.5-RMSProp, COURSERA: Neural networks for machine learning, University of Toronto, Tech. Rep. 2012.

27 S. K. Kalva and M. Pramanik, "Experimental validation of tangential resolution improvement in photoacoustic tomography using modified delay-and-sum reconstruction algorithm," *J. Biomed. Opt.* **21**(8), 086011 (2016).

28 American National Standard for Safe Use of Lasers ANSI Z136.1-2000.

29 Y. Lou et al., "Generation of anatomically realistic numerical phantoms for optoacoustic breast imaging," *Proc. SPIE* **9708**, 970840 (2016).

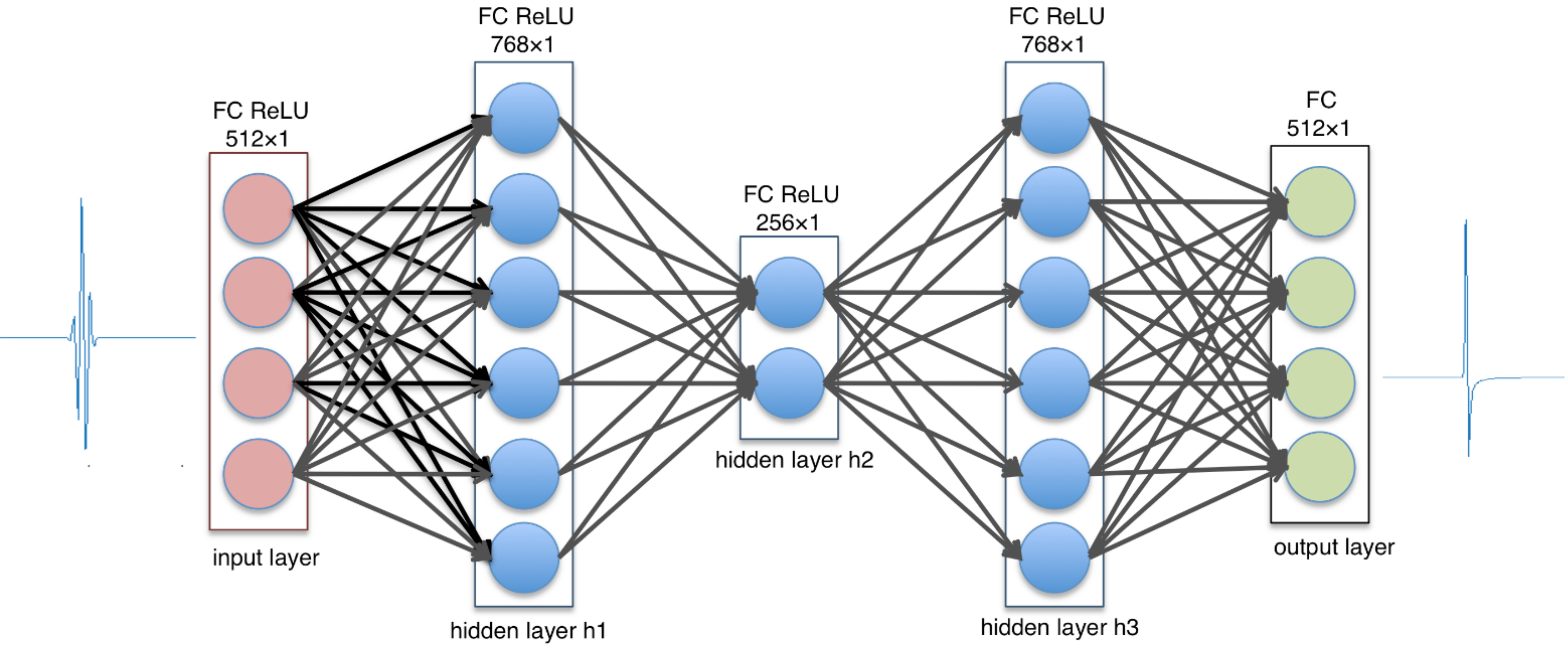
30 B. E. Treeby and B. T. Cox, “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wavefields,” *J. Biomed. Opt.* **15**(2), 021314 (2010).

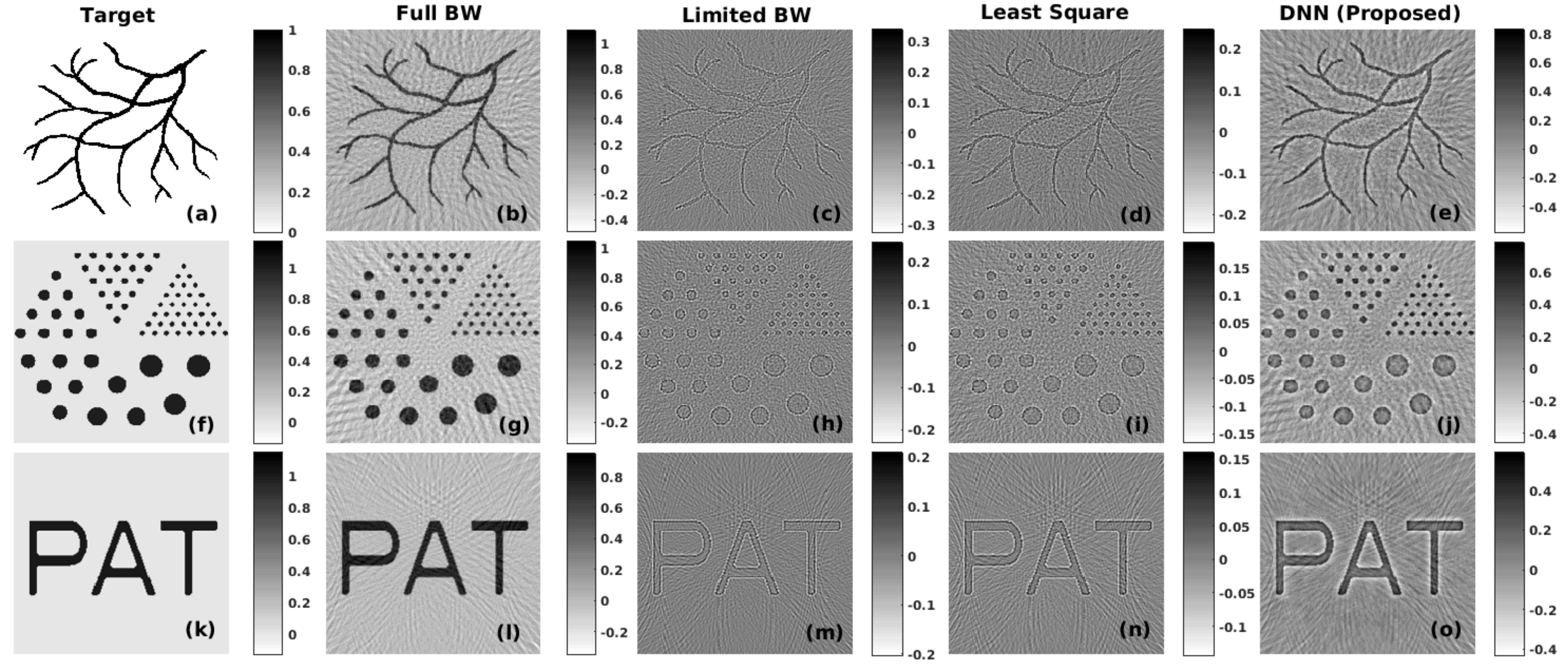
31 L. Jiang et al., “Self-paced learning with diversity,” *Advances in Neural Information Processing Systems*, 2078–2086 (2014).

32 Web-link: https://sites.google.com/site/sercmig/home/dnnbwe (accessed on: July 25, 2017).

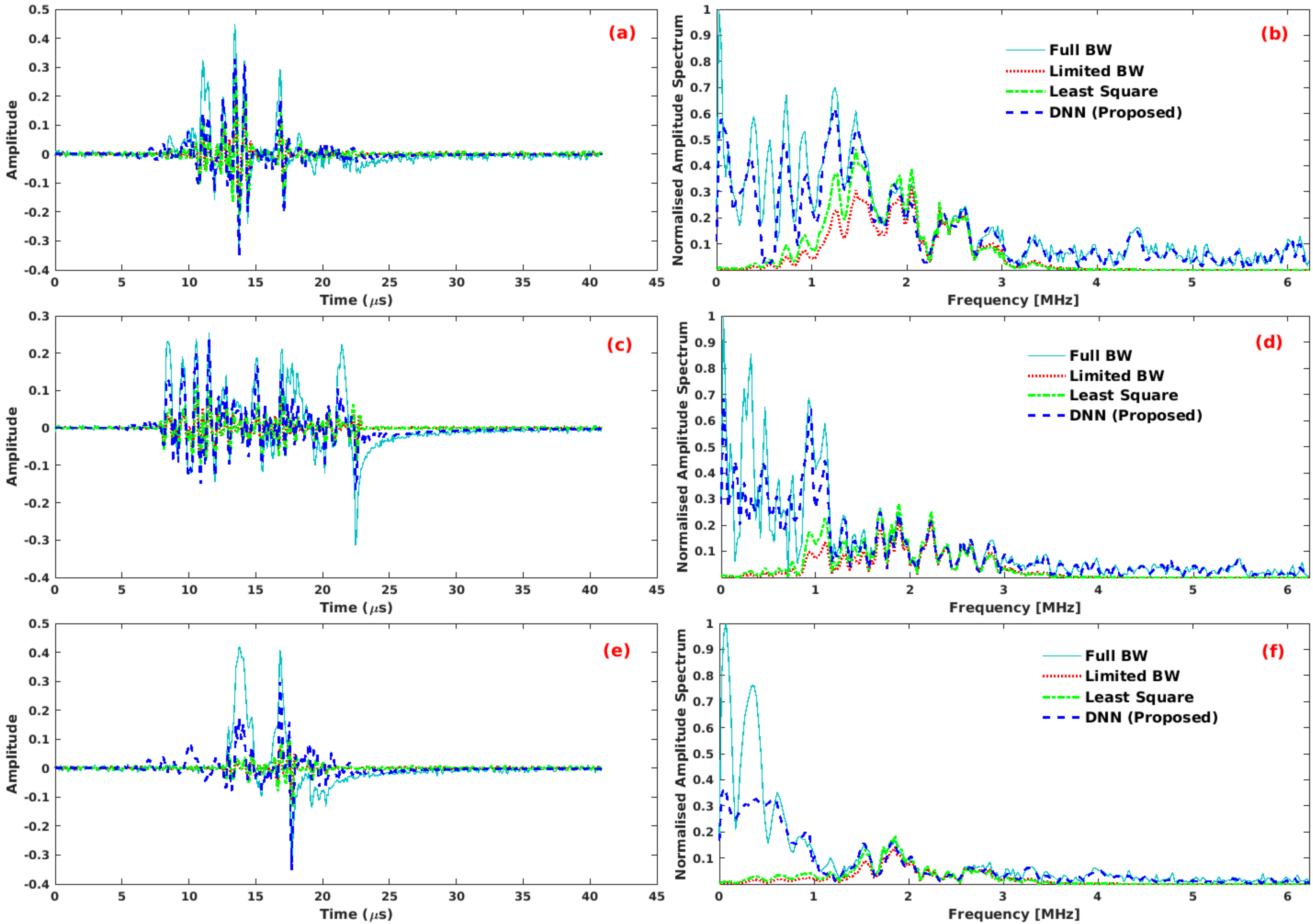
List of Figures

List of Tables

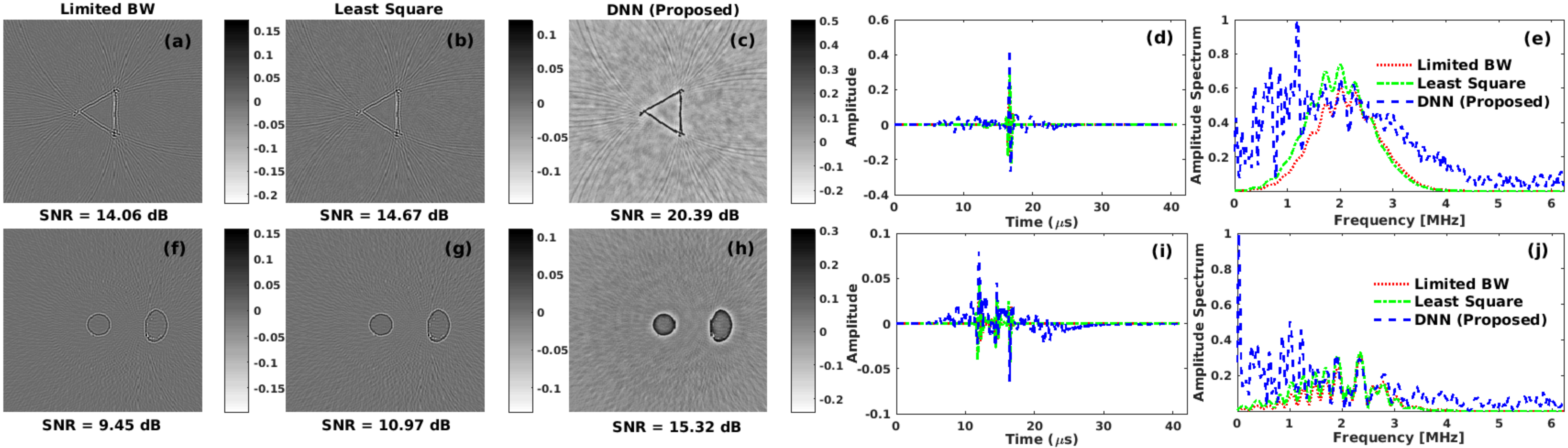

FC ReLU
512×1
input layer
FC ReLU
768×1
hidden layer h1
FC ReLU
256×1
hidden layer h2
FC ReLU
768×1
hidden layer h3
FC
512×1
output layer



Target
Full BW
Limited BW
Least Square
DNN (Proposed)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
PAT



(a)
Amplitude
Time (μs)
(b)
Normalised Amplitude Spectrum
Frequency [MHz]
Full BW
Limited BW
Least Square
DNN (Proposed)
(c)
Amplitude
Time (μs)
(d)
Normalised Amplitude Spectrum
Frequency [MHz]
Full BW
Limited BW
Least Square
DNN (Proposed)
(e)
Amplitude
Time (μs)
(f)
Normalised Amplitude Spectrum
Frequency [MHz]
Full BW
Limited BW
Least Square
DNN (Proposed)



Limited BW
Least Square
DNN (Proposed)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
SNR = 14.06 dB
SNR = 14.67 dB
SNR = 20.39 dB
SNR = 9.45 dB
SNR = 10.97 dB
SNR = 15.32 dB
Amplitude
Time (μs)
Amplitude Spectrum
Frequency [MHz]
Limited BW
Least Square
DNN (Proposed)