Ultrasonics 103 (2020) 106069

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras

A unified deep network for beamforming and speckle reduction in plane wave imaging: A simulation study

Etai Mor[a,b,*], Aharon Bar-Hillel[b]

[a] Department of Non Destructive Testing, Soreq Nuclear Research Center, Yavne 81800, Israel
[b] Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

ARTICLE INFO

Keywords:
Plane wave imaging
Ultrasound beamforming
Speckle reduction
Deep learning
Neural networks

ABSTRACT

Plane Wave Imaging is a fast imaging method used in ultrasound, which allows a high frame rate, but with compromised image quality when a single wave is used. In this work a learning-based approach was used to obtain improved image quality. The entire process of beamforming and speckle reduction was embedded in a single deep convolutional network, and trained with two types of simulated data. The network architecture was designed based on traditional physical considerations of the ultrasonic image formation pipe. As such, it includes beamforming with spatial matched filters, envelope detection, and a speckle reduction stage done in log-signal representation, with all stages containing trainable parameters. The approach was tested on the publicly available PICMUS datasets, achieving axial and lateral full-width-half-maximum (FWHM) resolution values of 0.22 mm and 0.35 mm respectively, and a Contrast to Noise Ratio (CNR) metric of 16.75 on the experimental datasets.

1. Introduction

Plane-Wave Imaging (PWI) is an emerging ultrasonic imaging modality which enables fast acquisition of the entire medium by emitting few non-focused, tilted, plane waves from a transducer array. The waves reflected from in-homogeneities in the inspected medium are separately recorded by each transducer. Subsequently, an image of the medium is produced by a beamforming process which combines the received signals in order to allow localization of the scattering structures with improved Signal to Noise Ratio (SNR). Since each tilted plane wave covers the entire medium, a coarse image can be produced from each plane wave and by further coherently combing the separated images, a fine image of the medium can be produced [1,2]. Generally, as the number of emitted plane waves is increased, the quality of the produced image is improved; however, the achievable frame rate is reduced. Therefore, there is a high interest in advanced beamforming methods which are capable to produce high quality images from few plane wave shots [3].

The most common beamforming method is the Delay and Sum (DAS) method [4]. In this method, a weighted sum of the time-delayed Radio-Frequency (RF) signals is constructed, with the delay based solely on the position of the transducer elements and the imaged point in the medium. While the DAS method is fairly simple, and can produce images in real time, it suffers from low lateral resolution and contrast quality, especially when using a small number of emitted plane waves [5]. In the past decade many advanced beamforming methods have been proposed. Examples are methods based on minimum variance [6], phase coherence [7–9], Delay-Multiply-And-Sum methods [10,11], Spatial Matched Filters (SMF) [12–14] and sparsity based methods [15,16]. Historically, these methods concentrate on enhancing the received signals from the focused image point, and reducing off-focus noise. While this improves the resolution and usually increases the contrast of the reconstructed images, it does not reduce the granular speckle pattern, common in medical ultrasonic images.

Speckle is caused by scattering sources which are smaller than the resolution of the imaging system and are not typically isolated. These scatterers cause constructive and destructive interferences of the backscattered waves which are expressed as a granular speckle pattern in the beamformed image. The speckle can be used to determine scattering properties of the tissue [17] and is also used in tracking of small displacements in the bloodstream [18]. However, speckle will also tend to reduce the contrast of reconstructed images and obscure macroscopic properties of the studied biological tissues, and is therefore generally treated as undesirable [19].

Speckle reduction techniques have been traditionally implemented by post-filtering of the already beamformed image. Common techniques include anisotropic diffusion methods [20–22], methods based on wavelet denoising [23,24], and non-local means methods [25–27]. These

* Corresponding author at: Department of Non Destructive Testing, Soreq Nuclear Research Center, Yavne 81800, Israel.
E-mail addresses: moret@post.bgu.ac.il (E. Mor), barhille@bgu.ac.il (A. Bar-Hillel).

https://doi.org/10.1016/j.ultras.2020.106069
Received 7 June 2019; Received in revised form 5 January 2020; Accepted 7 January 2020
Available online 23 January 2020
0041-624X/

methods treat speckle as unwanted noise in the beamformed image and try to remove it while maintaining the boundaries between relevant tissues. The main drawback of these post-processing methods is that they treat speckle filtering as a process which is independent from the beamforming process, and therefore neglect RF channel data which is lost during the beamforming process.

Recently Deep Neural Networks (DNN) has arisen as a promising framework, providing state-of-the-art performance in a variety of signal processing and machine learning tasks [28]. These methods are improvements of classical neural networks, including networks with more layers, permitting higher levels of abstraction, and improved techniques for network optimization and regularization. They have been found highly successful in image analysis tasks such as classification [29,30] and segmentation [31–33]. Recently, these methods were applied to image processing problems such as de-noising [34,35], de-convolution [36], super resolution [37,38] and medical image reconstruction [39], and there are indications of improvements over state-of-the-art methods.

Following the success of deep learning in these domains, the ultrasonic community proposed DNN methods for ultrasonic beamforming and related image processing tasks. In [40] deep learning was applied to the RF channel data in order to compress and recover ultrasound images. In [41] a beamforming network which operates on each frequency sub-band was proposed in order to suppress off-axis scattering. For limited angle PWI beamforming, [42] proposed a conventional deep Convolutional Neural Network (CNN) which operates on a limited angle RF channel data in order to obtain high quality images. Training was based on empirical images obtained from the full angle dataset.

DNN methods have also been proposed for speckle reduction of ultrasonic images. In [43] a CNN was proposed to approximate conventional despeckling methods in order to substantially reduce inference time. In addition, a network was trained to obtain CT-like quality by training on simulated ultrasound data based on real human body CT scans. In [44] despeckling was achieved by a DNN based on a Generative Adversarial Network (GAN), trained with patches of high quality and low quality ultrasound images. A GAN based network was also used in [45] in order to improve the runtime of high-computational methods such as [27].

In this work the entire process of PWI from a single angle is embedded in a single CNN, including both the beamforming and speckle reduction stages. The network follows a classical ultrasonic reconstruction pipeline of beamforming, envelope detection and speckle suppression. However, in contrast to the classical reconstruction process, details of the signal-to-image inference are learned from data, with all network stages including trainable parameters. These parameters are learned from a large set of simulated ultrasonic data, using several distributions of interest. This allows searching and optimization in a significant, large design space around the basic techniques.

The embedding of PWI into a convolutional network requires several non-trivial architectural choices. First, the beamforming stage is approached using an SMF approach, and is reduced to convolutional computations using a set of depth-specific filters learned from data. After beamforming, up-sampling and envelope detection with trainable filters is performed. Then, the image is logarithmically compressed to account for the large dynamic range of ultrasound signals and to match the log-domain in which the images are usually displayed. The log operation also turns the multiplicative speckle pattern to an additive one, which is easier to model [19]. For speckle removal, a multi-scale trainable architecture termed UNET [33] is adapted, which is able to approach and infer signal statistics in large arbitrary image regions.

Recently in [46], a different approach for embedding the entire inference (RF-channel data to de-speckled image) into a single DNN was proposed. While our network follows classical elements of the ultrasonic reconstruction pipeline, the network in [46] uses a conventional CNN structure. Neurons in such a network have a small limited receptive field, as opposed to the learnable SMF proposed by our approach, and therefore it cannot learn the appropriate time delays in order to focus the array to each imaged point. Hence, the appropriate time delayed RF signals are calculated in advance and feed into the network. This results in a very large input of size $N_z \times N_x \times N_{elm}$ where $N_z \times N_x$ represents the number of imaged points and N_{elm} is the number of elements in the array. Such a large input is only feasible for imaging small areas ($N_z = N_x = 64$ in [46]); however it is not appropriate for PWI applications, in which one usually images a large area at once using a large transducer array.

In Section 2 the proposed network model is presented. Section 3 discusses the training procedure of the network along with details of the simulated dataset used. In Section IV, the network is evaluated using simulated and real data, including experimental phantom and in-vivo data publically available by the PICMUS challenge [47]. The empirical results indicate that the learned network provides contrast and resolution significantly superior with respect to traditional DAS algorithm followed by common speckle reduction methods.

2. Network design

Our network architecture is based on a CNN receiving as input a matrix of $N_t \times N_{\text{elm}}$, representing the channel RF signals acquired by N_{elm} transducers, sampled at N_t time steps. It outputs a reconstructed image of size $N_z \times N_x$ ($N_t = 2,000$, $N_{\text{elm}} = 128$, $N_z = 480$ *and* $N_x = 384$ in the experiments – see Section 4). The proposed network, presented in Fig. 1, consists of two main parts: (1) a beamforming network based on an SMF architecture, discussed in Section 2.1; (2) a speckle reduction module, based on a residual UNET architecture [33] (described in Section 2.2). The latter is applied after envelope detection, transforming to log space and a 2-layer convolutional sub-network.

2.1. Beamforming by a spatial matched filters network

The beamforming network is based on a matched filter method for enhancing the SNR of the received RF signals [12–14]. Generally, SMF calculates the beamformed image $y(z, x)$ at location (z, x) by preforming correlations between the received RF signals $s(t_i, x_r)$ (sampled at $\{t_1, \cdots, t_{N_t}\}$ time samples and $\{x_1, \cdots, x_{N_{\text{elm}}}\}$ transducer locations) and a predefined filter $G(t_i, x_r; z, x)$ which represents the RF signals obtained from a point reflector located at the imaging point(z, x):

$$y(z, x) = \sum_{i=1}^{N_t} \sum_{r=1}^{N_{\text{elm}}} s(t_i, x_r) G(t_i, x_r; z, x). \tag{1}$$

For PWI, we may assume that for a given depth z, the filters are laterally invariant, i.e. there are x-independent filters $F(t, l; z)$ such that $G(t_i, x_r; z, x) = F(t_i, x_r - x; z)$ and equation (1) can be written as a convolution of the form

$$y(z, x) = \sum_{i=1}^{N_t} \sum_{r=1}^{N_{\text{elm}}} s(t_i, x_r) F(t_i, x_r - x; z). \tag{2}$$

While in classical implementations of SMF [12–14], $G(t_i, x_r; z, x)$ is based solely on a physical description of the spatial impulse response, in this work the matched filters are learned from the training data. Moreover, since Eq. (2) defines a convolution, for a fixed depth z, it can be implemented using a standard convolutional layer with a single filter. In order to use a finite set of filters, the depth range is quantized into a set $\{d_1, \cdots, d_{N_z}\}$ of N_z distinct depth values, where the number of filters N_z is application dependent and defines the desired axial output resolution. The size of the filters ($H_j \times D_j$ for $j = 1, .., N_z$ in Fig. 1) is depth dependent and is set to capture the support of the theoretical spatial impulse response of a scatterer at depth d_j. The output of each filter is a single line of size $(1 \times N_{\text{elm}})$. These lines are merged together to form an image of size $(N_z \times N_{\text{elm}})$.

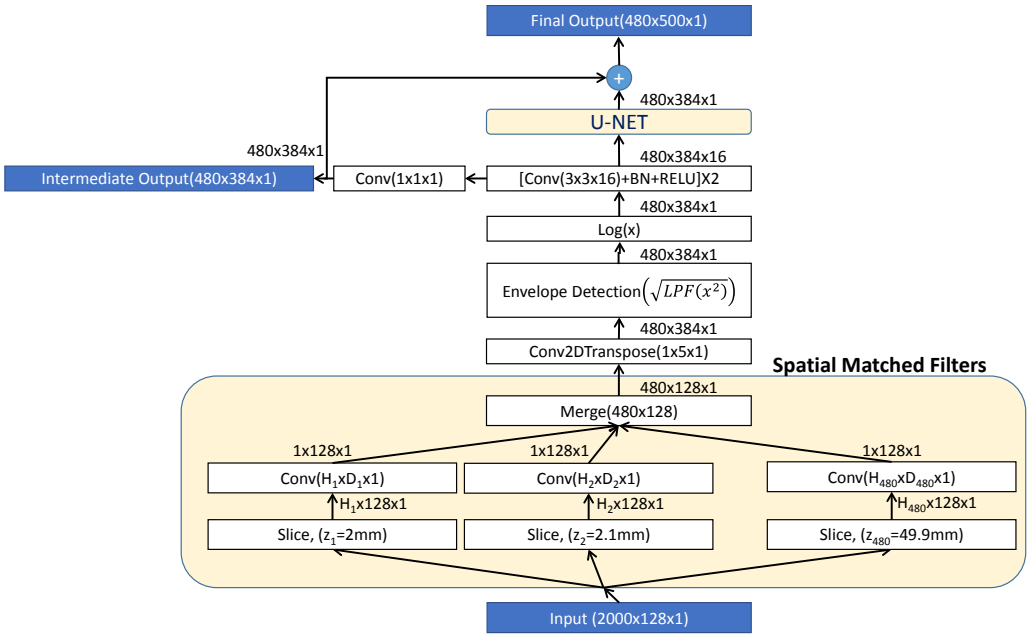

Fig. 1. High level architecture of the proposed deep convolutional network. The beamforming subnet is based on a set of depth specific two dimensional convolutions whose output is combined into a single image, followed by up-sampling (by a transpose convolutional layer), envelope detection, log transformation, and a small two-layer convolutional network resulting in 16 filtered channels. These channels are the source for two outputs. The first is the intermediate output which is achieved by a 1 X 1 convolutional filter and is aimed at identifying isolated scatters. The second is the final-output which is performed by a UNET network, described in detail in Fig. 2, and is aimed for speckle reduction.

Following the depth-dependent filtering, the resulting map is up-sampled in the lateral direction to the required number of pixels ($N_z \times N_x$) using a learnable transpose convolutional layer initialized to a bilinear interpolator [48]. Then, oscillations caused by the wave nature of the beamformed signals are removed. This is done by a square-law envelope detector [49,50], applied to every column x, of the form

$$v(d_n, x) = \sqrt{\sum_{k=-L}^{L} h(L+k) y^2(d_n + k, x)} \quad 0 \le d_n < d_{N_z} \tag{3}$$

where $y(z, x)$ is the up-sampled input image and $h \in R^{2L+1}$ is a low pass filter. The filter is learned, but it is initialized to a hamming low-pass filter with a cutoff wave-number of f_c/c, where f_c is the center frequency of the signals and c is the material velocity.

Following the envelope detection, a log transformation is applied to the image in order to compress the large dynamic range of ultrasound signals and to match the log-domain in which the ultrasonic images are usually displayed. Furthermore, the log operation also transforms the multiplicative nature of the speckle pattern into an additive one [19], which is easier to remove (Section 2.2).

Finally, the log-compressed image is further processed by two convolutional layers, each composed of 16 filters of size 3 X 3, followed by Batch Normalization (BN) [51] and a Rectified Linear Unit (ReLU) [52]. This result in 16 filtered maps, which are combined by a 1 X 1 convolutional filter into a single beamformed image (presented as an intermediate output in Fig. 1).

2.2. Speckle reduction by a residual UNET architecture

The network presented in Section 2.1 was designed in order to infer a beamformed image representing localized scatterers, indicating changes in the mediums impedance. However, it was not designed to reduce speckle which is usually of a larger spatial extent. In order to reduce such speckle pattern, a modified version of the UNET architecture, originally proposed for medical image segmentation [33], is applied. For this network, we take as input the 16 upper channels calculated by the network presented in Section 2.1 (see Fig. 1). Then, instead of directly learning the transformation between these maps and a true clean image, which may lead to slow convergence [53], the residual between the beamformed image (intermediate output in Fig. 1 which contains speckle noise) and the clean image is learned. This residual amounts to the (negative of the) speckle pattern, and the network learns this pattern by training a UNET [33] structure, presented in Fig. 2.

The UNET network is a type of Encoder-Decoder network with skip connections (presented by horizontal blue lines in Fig. 2). The blue boxes in Fig. 2 represent feature maps produced in the encoder stage (left side of the U structure). These are formed by repeated application of 3×3 convolutional filters followed by Batch Normalization (BN) [51] and a ReLU nonlinear function [52] (represented by horizontal red arrows). Processing is sequentially done in multiple image scales (spatial resolutions). After each encoder stage, feature maps are transferred to a coarser spatial scale by max pooling and down sampling (the azure arrows in Fig. 2). The final feature maps at each image resolution are then integrated by the decoder (right side of the U structure). At decoder scale i, an initial feature map is created by concatenation of the feature maps from scale i of the encoder and decoder features maps of the higher resolution i + 1. The latter are up sampled to reach scale i and are going through a 1×1 convolution to reduce the number of maps by a factor of 2. At each scale, the decoder processes the maps with several 3×3 convolutional maps, before up-sampling them to the next scale. The final feature maps are at the resolution of the original input channels, yet they include feature maps with neurons of very large receptive field, capable of multiple scale analysis. A final 1×1 convolution layer is then used to output the desired residue image.

Such a network provides good spatial integration capabilities by including features from multiple image resolutions interacting in a single final layer. Since each output pixel should describe the reflectivity in a cluster of unknown size around it (see Section 3.2), such multiple-scale integration of spatial information is required.

3. Network training

Two datasets were generated for network training, each containing 10,000 simulations. The first (details in Section 3.1), includes simple simulations composed of highly separated point reflectors (scatterers). It is used for training of the beamforming sub-network weights. The second training set (details in Section 3.2) is based on a more involved simulation mimicking realistic scenarios, where different regions of speckle intensities exist. End-to-end learning of the SMF weights using only the second data set, which does not contain isolated scatterers, poses a difficult learning problem for the deep network, and our preliminary experiments exhibited learning failures in this setting. The

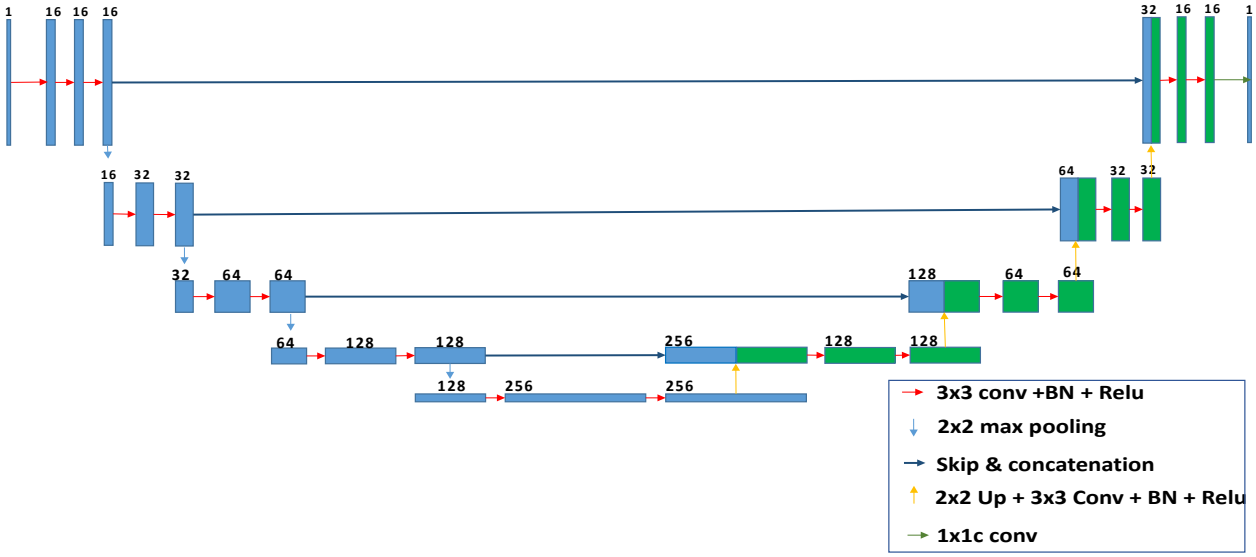

Fig. 2. The UNET [22] architecture utilized for modeling the speckle pattern. The network enables reasoning over image areas of multiple scales, required for inferring the cluster to which a pixel belongs. Such inference is the key for averaging over possibly-large unknown image areas, as required for speckle modeling and reduction.

Table 1
Transducer Parameters.

Parameter	Value
Pitch	0.30 mm
Element width	0.27 mm
Element height	5 mm
Elevation focus	20 mm
Number of elements	128
Aperture width	38.4 mm
Transmit frequency	5.208 MHz
Sampling frequency	20.832 MHz
Pulse bandwidth	67%
Excitation	2.5 cycles

difficulty is due to the indirect inference involved, and the large number of parameters in the SMF filters, which are of a large spatial extent. We therefore used a two-stage learning phase in which the weights learned from the first dataset were used as initialization for training with the second set.

All simulation data was created using the Field-II ultrasound simulation program [54]. The material velocity was set to 1.54[mm/μsec] and all scatterers where positioned between lateral position $x \in [-25, 25]$mm and depth$z \in [2, 50]$mm. In all simulations the transducer parameters defined by the PICMUS challenge [47] were used, presented in Table 1.

Training was done with the ADAM gradient-based optimization method [55] with a mini batch size of 10 images and a learning rate of 0.001. Each dataset was divided to 8,000 training simulations and 2,000 simulations in a validation set. Training proceeded until the error on the validation set did not improve for 4 epochs. Networks were implemented using the Keras [56] software with a Tensorflow [57] backend and took about 8 hours on an NVidia 1080Ti GPU.

3.1. Beamforming training using sparse scatterers simulations

For training the beamforming network 10,000 simulations were performed, each with 1,000 scatterers of a constant reflectivity. Each simulation generated a measurement matrix of size $N_t \times N_{\text{elm}} = 2{,}000 \times 128$ representing the RF signals received from a single plane wave at an angle of 0^o. Additionally, 0.1 mm resolution scatterers-images of size$N_z \times N_x = 480 \times 384$ were produced, defining the ground truth output. Scatterers were positioned in the region $x \in [-19.5 \text{ mm}, 19.5 \text{ mm}]$ and $z \in [2 \text{ mm}, 50 \text{ mm}]$ relative to the transducer, which was centered in $(x = 0, z = 0)$ and faced the positive z axis direction (z axis points downwards in all subsequent figures)[1]. In Fig. 3 a sample image of the envelope-RF measurement matrix (Fig. 3(a)) is shown, along with its respective ground truth image (Fig. 3(b)).

Training was performed with a L_2 loss between the estimated images (intermediate output in Fig. 1) and the ground truth images. L_1 regularization on the matched filters weights was added since these are expected to be highly sparse. The low-pass filter of the envelope detector was initialized to a hamming filter while all other weights were randomly initialized using the Xavier uniform initialization scheme [58].

The learned filters, represented by $F(t, l; z)$ in equation (2), for $z = 10, 20, 30$ *and* 40mmare presented in Fig. 4. On top of the filters, red lines present the delay times $dT(l; z)$ which were computed according to the time it takes for a plane wave to reach a scatterer positioned at depth z and laterally diverted by $l = x - x_n$ from transducer elements, positioned at $\{x_1, \cdots, x_{N_{\text{elm}}}\}$:

$$dT(l; z) = \frac{z + \sqrt{l^2 + z^2}}{c}, \tag{4}$$

where c is the medium sound velocity. One can observe that the learned filters are indeed very sparse and follow adequately the spatial spread of the beam (4). Moreover, the filters reflect the temporal oscillatory nature of the received RF signals related to the transducers bandwidth and central frequency. Finally, it is clear that the filters also perform channel apodization, by giving less weight to channels farther from the center transducer element. The effective channel support becomes smaller for smaller depths. Such filters, which describe the impact of a single isolated scatterer, allow precise localization with improved *SNR*.

In Fig. 5 a reconstruction result example from the validation set is presented (utilizing the trained filters presented in Fig. 4). One can observe the improved resolution of the proposed method (Fig. 5(c)), allowing separation of close scatterers which is not possible using the

[1] While the simulations contained scatterers in a large region of 50 mm × 50 mm, the reconstruction was done only for the 48 mm × 38.4 mm area in front of the transducer in order to comply with the reconstruction region of the PICMUS datasets (see experiments in Section 4).

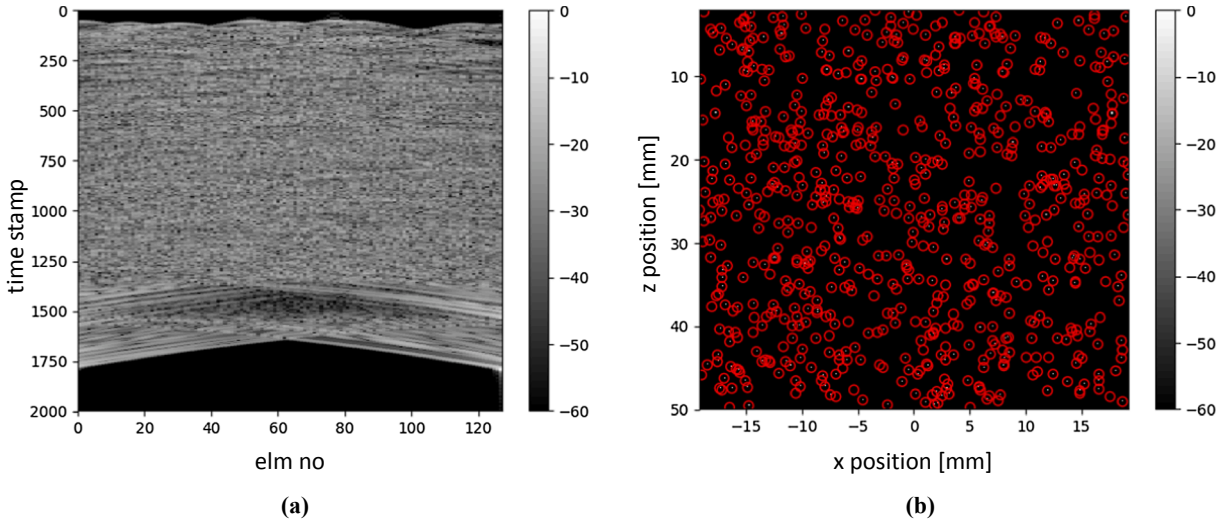

Fig. 3. Simulation of sparsely distributed scatterers. (a) The 2,000 × 128 envelope-RF measurements matrix. (b) Ground truth scatterers-image, the scatterers are marked by red circles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

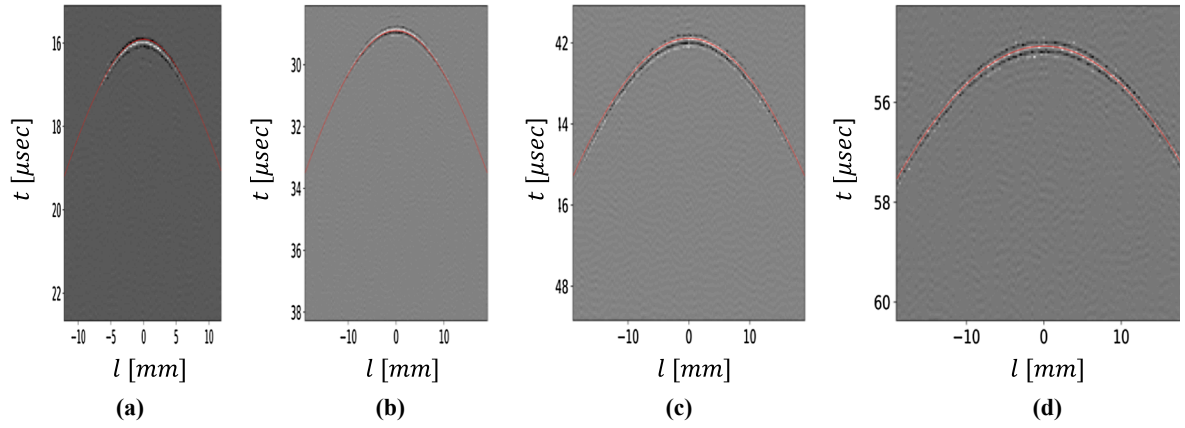

Fig. 4. Learned spatial matched filters at depths 10 mm (a), 20 mm (b) 30 mm (c) and 40 mm (d). Delay times of a single scatterer are marked in red. These learned filters, which are represented by $F(t, l; z)$ in equation (2), are very similar to the impulse responses of single scatterers positioned at the different depths. One can see that 1) The filters encode the time delays of a single scatterer. 2) For small depths arcs have smaller effective support, reflecting the fact that transducers elements facing the scatterer are closer to it, and hence they have larger weights. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

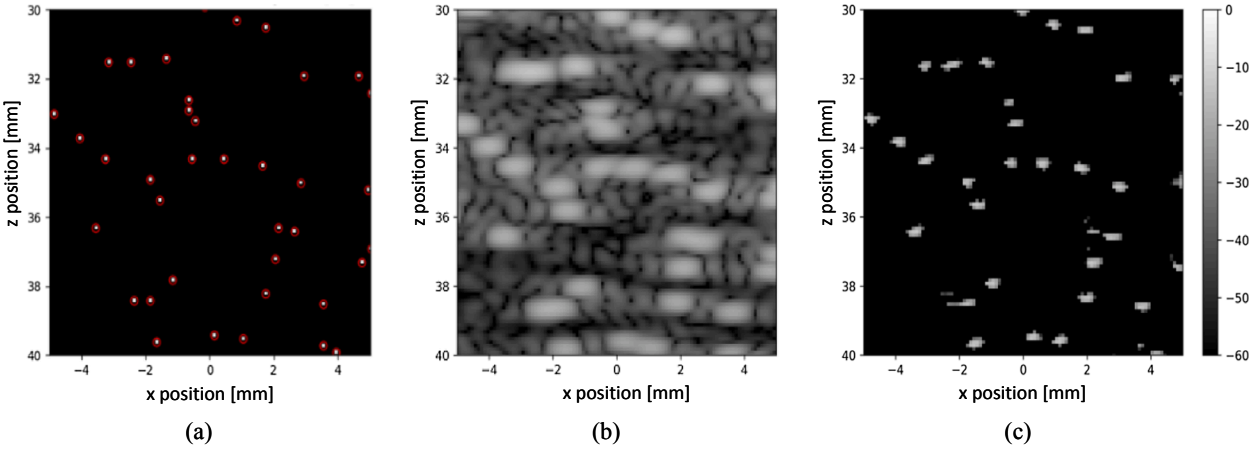

Fig. 5. Reconstruction results from a simulation of sparsely-distributed scatterers. (a) Ground truth scatterers - this is a zoom-in on a sub-region in a validation example. (b) DAS reconstruction. (c) Proposed network – Intermediate output.

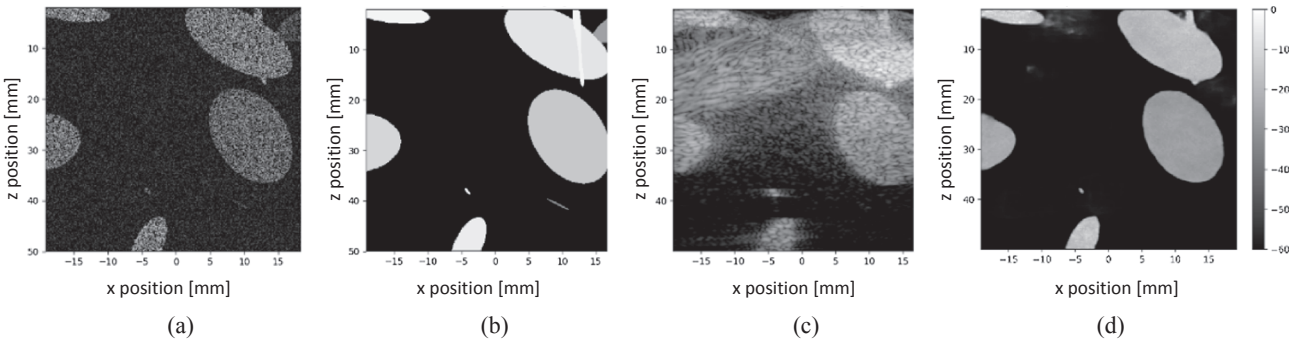

Fig. 6. Fully resolved speckle simulation containing different ellipses with different standard deviation of scatterer reflectivity, along with reconstruction results. (a) Image of the densely-distributed scatterers. (b) Ground truth clusters presenting the standard deviation of scatterer reflectivity. (c) DAS reconstruction. (d) Reconstruction using the proposed network (final output).

DAS method (Fig. 5(b)). However, for medical ultrasound applications the scatterers are not typically isolated and their density is causing the speckle pattern in the beamformed image. In order to deal with this speckle noise, the whole network (presented in Fig. 1) has to be trained on a representative database containing densely-distributed scatterers.

3.2. Training utilizing fully resolved speckle simulations

10,000 phantoms were synthesized for full network training, containing different echogenicity regions of the imaged medium. 200,000 scatterers, positioned in a 2D plane facing the transducer (x-z plane), were sampled in each phantom in order to generate images with a fully developed speckle, in a similar manner to [59]. While this neglects the effect of scatterers positioned outside the x-z plane, it enabled us to generate a larger set of simulated phantoms (10,000) in a reasonable execution time. This amount of phantoms was needed in order for the network to generalize well on simulation and experimental phantoms (see Section 4).

In each phantom, we generated ten random two-dimensional elliptic clusters of different size, orientation and standard deviation of the scatterers. We have chosen ellipse shapes as our primitives since they serve as a good first-approximation to general convex shapes. It should be noted that while this may create some tendency to learning convex regions, the training data contains non-convex regions created by layered structure of multiple ellipses. The standard deviation of the scatterers' reflectivity represents the average echogenicity of the scatterers in a region [46]. Specifically, the inferred image resulting from a homogeneous region of tissue, of constant echogenicity, should produce a uniform response, as opposed to a speckle response, while preserving the structure and echogenicity of the region. The ellipse centers were drawn uniformly in the range $x \in [-25, 25]$ mm and $z \in [2, 50]$ mm. The ellipse semi-axes and orientations were drawn uniformly with $r \in [0, 10]$ mm and$\phi \in [0, \pi]$. The standard deviation of the scatterer reflectivity was drawn uniformly in a log scale of $[-2, 2]$ i.e. the standard deviation of each ellipsoid was set in the range$[10^{-2}, 10^{2}]$.

In Fig. 6(a)-(b) an image of densely-distributed scatterers is presented, along with its ground truth cluster image (taken from the validation set). The transducer and the transmission scheme along with the medium properties were all identical to the sparsely-distributed simulations (Section 3.1).

The network was trained by a combination of two losses. The first is a L_2 loss between the scatters reflectivity amplitude image (like Fig. 6(a)) and the intermediate output of the network. This loss preserves the ability of the beamforming network to recognize isolated scatters. The second is a L_1 loss, demanding that each output pixel will match the echogenicity (standard deviation of the scatterers' reflectivity) in its local cluster, i.e. it is trained to infer the speckle-free image as shown in Fig. 6(b). The training objective was to minimize an average of the two losses. Beamforming-network weights were initialized to the values learned using the separated point reflector dataset (Section 3.1), while the UNET network weights were randomly initialized using the Xavier uniform initialization scheme [58].

Reconstruction results of this simulation are presented in Fig. 6(c)–(d) for the DAS method[2] and our final network output respectively. One can observe the better contrast ratio and the smoother results of the proposed network. As will be discussed later, these smoother results do not come on the expense of the axial or lateral resolution of the output.

4. Results

In this section the network is quantitatively tested on external datasets publically available by the PICMUS challenge [47]. The RF signals in all simulations and experiments were acquired using the same transducer, with the parameters from Table 1 (This transducer was also used in the simulations for training our network). We compare our method to DAS as well as DAS followed by an influential speckle reduction technique: DPAD [21], OBNLM [26] and ADMSS [22]. Tuning parameters for the despeckling methods were set according to Table 1 in [22]. To better understand the contribution of the two network parts, we also tested a network including the UNET but without the beamforming network. In this variation, the UNET receives as input the log-compressed envelop of a single DAS beamformed image (This is contrary to the proposed network in which the network receives 16 learned maps as input). This UNET network was trained in a similar manner to the proposed network as described in Section 3.2. All images are presented using a constant 60 dB dynamic-range.

4.1. Resolution evaluation

For evaluating the resolution of the produced images, the PICMUS challenge provides two datasets obtained by scanning a target containing separated scatterers. The first, denoted by SR, is based on a noise free simulation of twenty separated scatterers. The second, denoted by ER, is based on a physically scanned in-vitro tissue mimicking phantom, containing seven nylon wires. Reconstruction results for the SR phantom by DAS, DPAD, OBNLM, ADMSS, DAS + UNET and the proposed method (intermediate and final outputs) are presented in Fig. 7.

Fig. 8 presents the axial and lateral response of the scatterer positioned at$z = 20mm$. It is clear that for this case, which does not contain any speckle, our proposed network achieved the best resolution results already in its intermediate output and these results were also retained in the final de-speckled output. The network achieved axial and lateral

[2] The DAS reconstruction software is freely available from the PICMUS site: https://www.creatis.insa-lyon.fr/Challenge/IEEE_IUS_2016

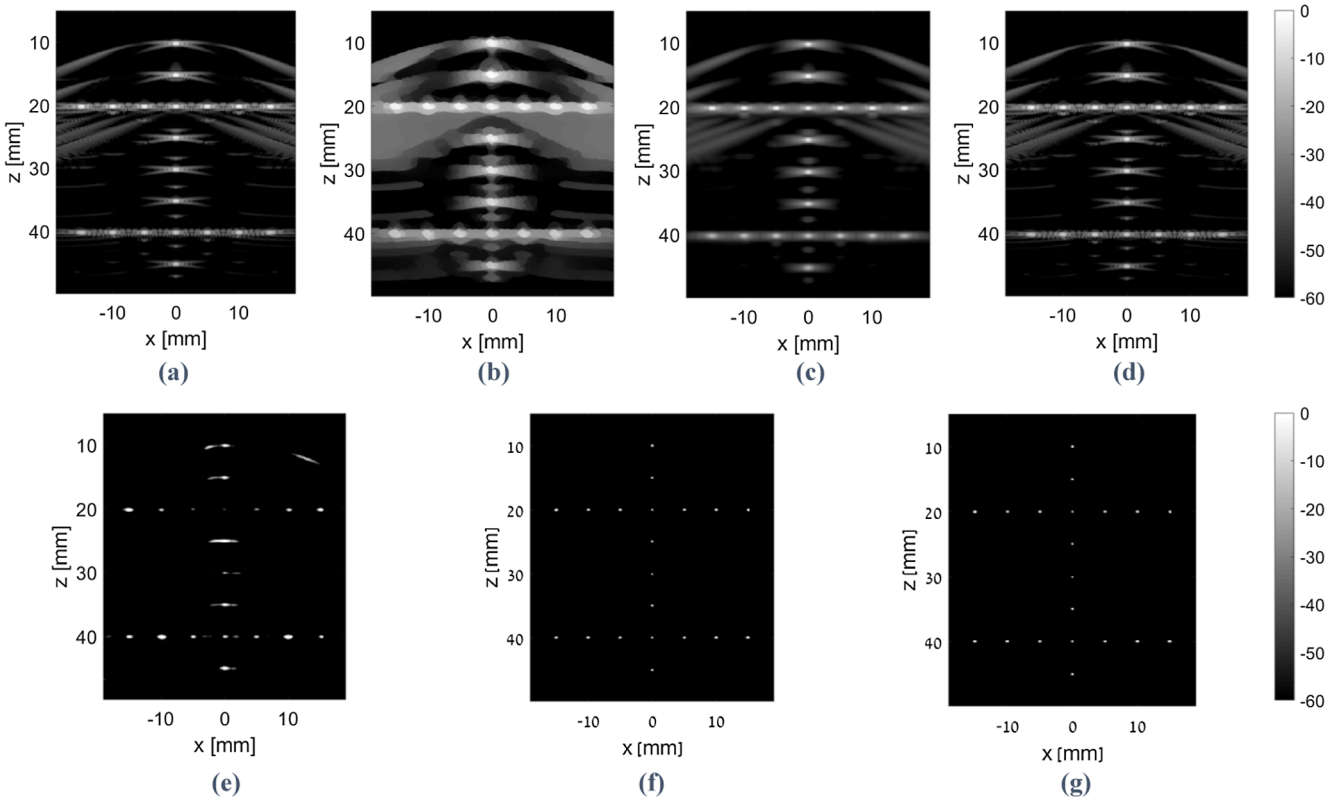

Fig. 7. Reconstruction results for the 20 point simulated target obtained by (a) DAS, (b) DPAD, (c) OBNMLF, (d) ADMSS, (e) DAS + UNET, (f) proposed network – intermediate output, and (g) proposed network – final output.

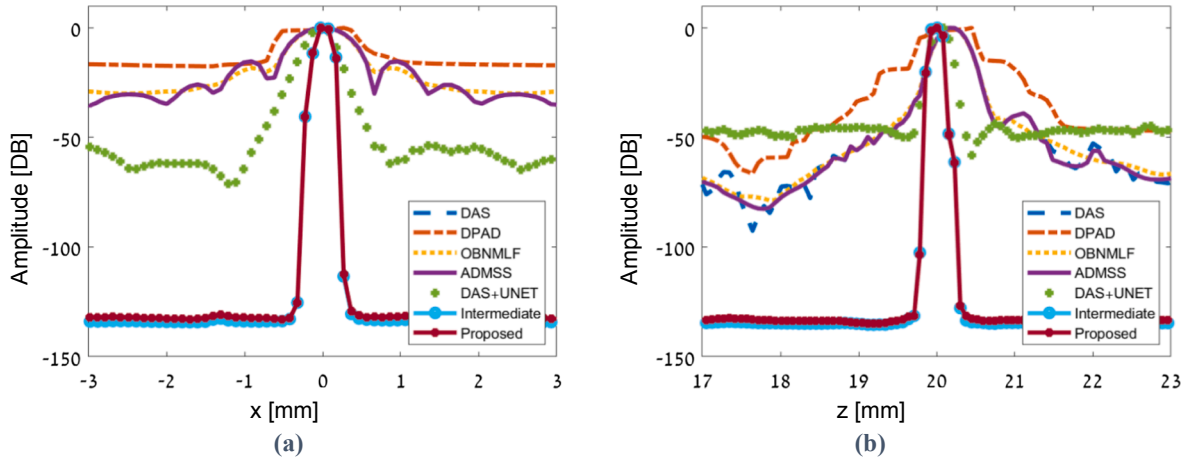

Fig. 8. (a) Axial and (b) lateral response of a simulated scatterer at depth 20 mm. Beam profiles are presented for the DAS, DAS followed by speckle reduction methods (DPAD, OBNMLF, ADMSS), DAS followed by a learnable UNET, and the proposed method (intermediate and final outputs).

Full-Width-Half-Maximum (FWHM) resolution indexes of 0.19 mm and 0.27 mm respectively (see Table 2- row no. 7). Good results were also achieved by the UNET + DAS method which improved over DAS and achieved axial and lateral FWHM values of 0.24 mm and 0.52 mm respectively. The OBNMLF and ADMSS despeckling methods obtained resolution results similar to DAS, while the DPAD method compromised the output resolution.

Fig. 9 presents the results for the ER phantom. Axial and lateral response of the scatterer positioned at 18.7 mm depth is shown in Fig. 10. For this case, in contrast to the SR dataset which did not contain speckle, the intermediate and final outputs of the network are significantly different. The intermediate output suppressed all scatterers except the most significant ones. While this resulted in good FWHM resolution indexes (0.27 mm and 0.41 mm for axial and lateral FWHM values respectively) it also removed significant tissue areas such as the high echogenicity circle present at($x \cong -10\ mm, z \cong 28\ mm$). Our final output, on the other hand, achieved the best FWHM resolution values (0.22 mm and 0.35 mm for axial and lateral FWHM values respectively) and was also able to suppress the speckle pattern while preserving different tissue areas like the circle at the middle-left of the image.

Regarding DAS followed by a despeckling method, it is seems that OBNMLF and ADMSS methods slightly reduced the speckle pattern while still preserving the DAS resolution (0.57 mm and 0.89 mm for axial and lateral FWHM values respectively). However DPAD, and DAS + UNET methods obtained an inferior resolution compared to the

Table 2
Evaluation of the methods on the PICMUS dataset in terms of axial and lateral FWHM resolution (A = Axial, L = Lateral) and CNR for the different datasets (SR, SC, ER, EC). Best performance metrics are marked in bold.

#	Method	SR-FWHM		SC	ER-FWHM		EC
		A	L	CNR	A	L	CNR
1	DAS	0.4	0.82	9.96	0.57	0.89	8.15
2	DAS + DPAD	0.81	1.25	14.61	0.87	1.3	12.9
3	DAS + OBNMLF	0.42	0.82	13.38	0.57	0.89	11.35
4	DAS + ADMSS	0.4	0.82	12.6	0.57	0.89	10.85
5	DAS + UNET	0.24	0.52	**27.1**	0.93	1	**17.3**
1	MV [60]	0.41	**0.12**	12.5	0.57	0.43	8.9
2	MV + DPAD	0.91	1.06	16.81	0.87	0.92	14.7
3	MV + OBNMLF	0.59	0.54	15.4	0.59	0.60	13.35
4	MV + ADMSS	0.41	**0.12**	14.12	0.57	0.44	12.9
9	Proposed –Intermediate output	0.19	0.27	8.81	0.27	0.41	4.55
10	Proposed –Final output	**0.19**	0.27	22.63	**0.22**	**0.35**	16.75

DAS method, and overly smoothed the image. It should also be noted that, for the output of the learned methods (DAS + UNET and the proposed network), there are some regions in the background which are not uniform and may be attributed to the network not generalizing well. This is probably due to training which was done solely on simulation datasets. In future research we intend to refine the training using more experimental data (see Section 5).

4.2. Contrast evaluation

Contrast of produced images in PICMUS was evaluated by two additional datasets denoted by SC and EC for simulated and experimental data respectively. These datasets were obtained by scanning of a target containing predefined scattering-free regions. For evaluating the contrast of the produced images, a Contrast to Noise Ratio (CNR) metric was calculated by

$$CNR = 20\log_{10}\left(\frac{|\mu_{in} - \mu_{out}|}{\sqrt{\sigma_{in}^2 + \sigma_{out}^2}}\right) \tag{5}$$

where μ_{in} and σ_{in} are the mean and standard deviation of the envelop signal intensities, measured in DB, inside the scattering-free region and similarly μ_{out} and σ_{out} are the mean and standard deviation outside the same region.

Results for the SC dataset are presented in Fig. 11. For this dataset, the DAS method achieved a CNR of 9.96 and the despeckling methods: DPAD, OBNMLF and ADMSS improved this to 14.61, 13.38 and 12.6 respectively. DAS + UNET achieved a CNR metric of 27.1 which is the best result for the SC dataset. The proposed method in its intermediate output achieved a poor CNR of only 8.81. This is due to the highly granular nature of the intermediate output which increases the standard deviations σ_{out} in (5). The final output, on the other hand, results in a clean image achieving a second best CNR result of 22.63 for the SC dataset.

Similar behavior is obtained for the experimental dataset (EC), presented in Fig. 12. Here the DAS method achieved a CNR of 8.15 and the despeckling methods: DPAD, OBNMLF and ADMSS improved this to 12.9, 11.35 and 10.85 respectively. DAS + UNET achieved a CNR metric of 17.3 which is the best result for the EC dataset. However, is seems that it overly smoothed the image suppressing important details such as the strong point reflector positioned around($x \cong 10$, $z \cong 37$). For this case the intermediate output suppressed almost all scatterers besides some at the right-center of the image (due to a hyperechoic area). Our final output, on the other hand, achieved a relatively smooth image attaining a CNR of 16.75 while still preserving fine details such as the strong scatterer at($x \cong 10$, $z \cong 37$).

A full summary of the results is presented in Table 2. In addition to DAS, the table also include evaluation of adaptive Minimum Variance (MV) method [60] followed by the DPAD, OBNMLF and ADMSS

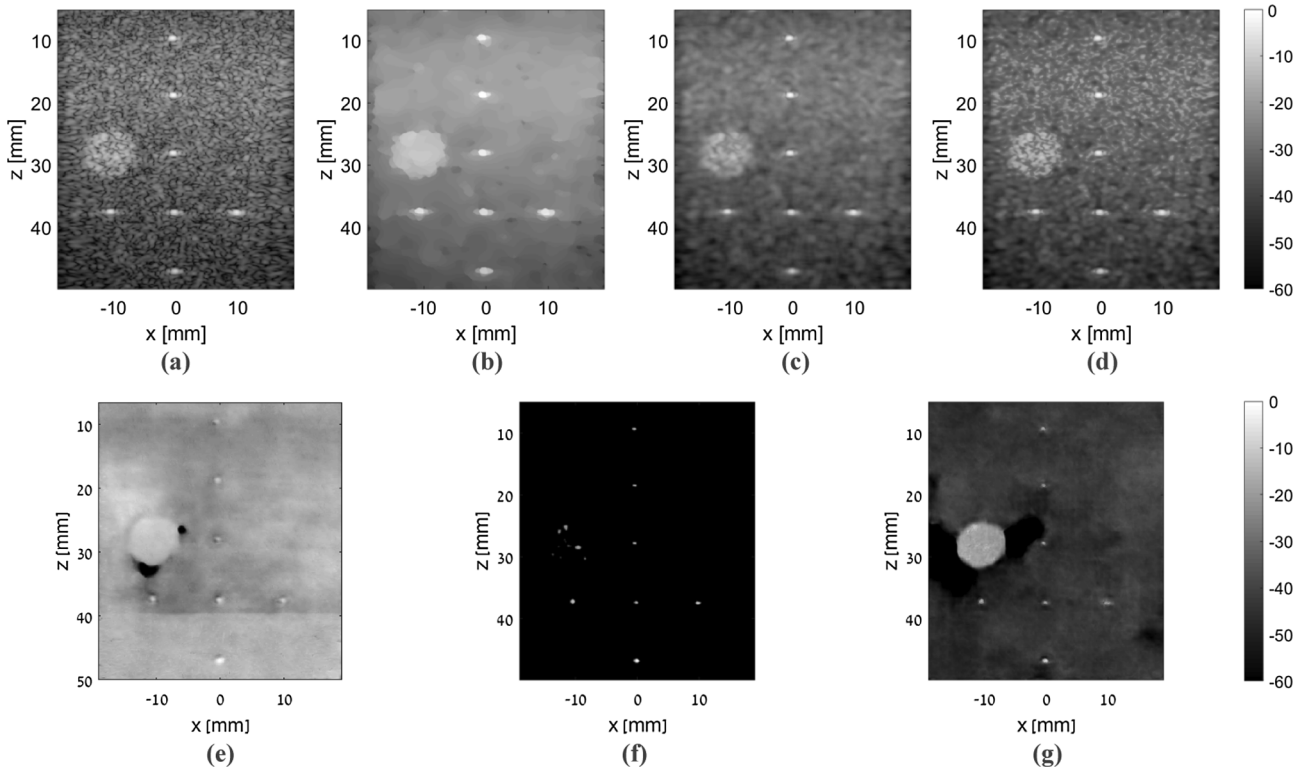

Fig. 9. Reconstruction results for the experimental nylon wires phantom. (a) DAS (b) DPAD (c) OBNMLF (d) ADMSS (e) DAS + UNET (f) Proposed network – intermediate output (g) Proposed network – final output.

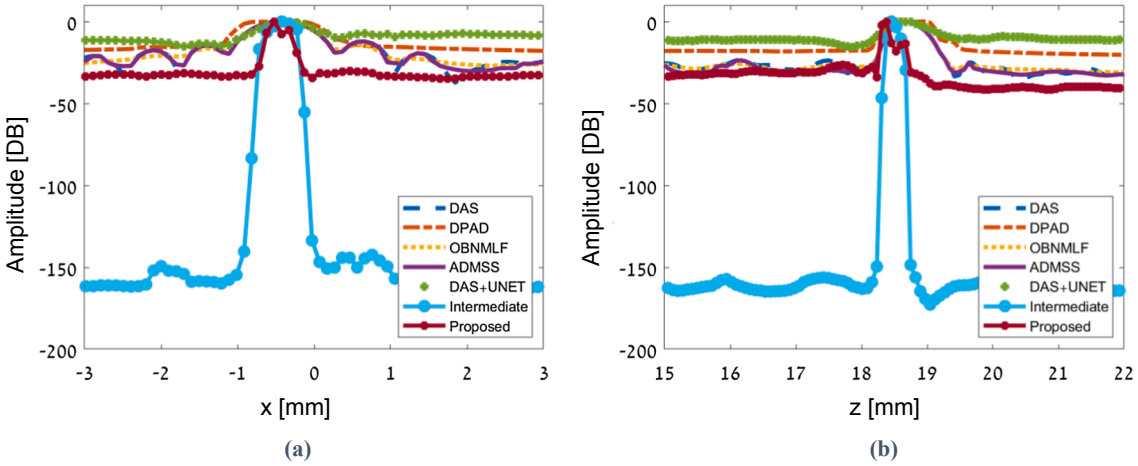

Fig. 10. (a) Axial and (b) lateral response for a nylon wire located at depth 18.7 mm. Beam profiles are presented for the DAS, DAS followed by speckle reduction methods (DPAD, OBNMLF, ADMSS), DAS followed by a learnable UNET, and the proposed method (intermediate and final outputs).

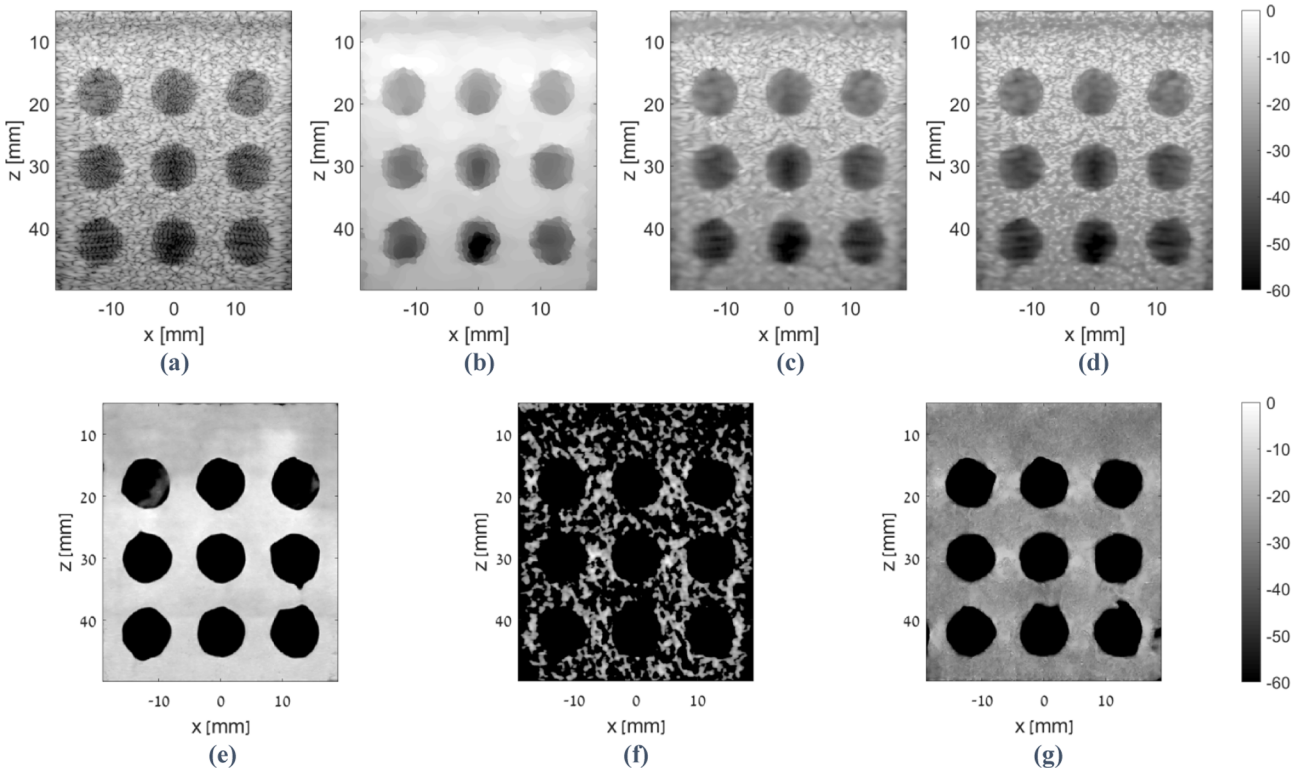

Fig. 11. Reconstruction results for the simulated contrast evaluation dataset. (a) DAS (b) DPAD (c) OBNMLF (d) ADMSS (e) DAS + UNET (f) Proposed network – intermediate output (g) Proposed network – final output.

despeckling methods. It can be seen that for the experimental datasets the proposed method achieved the best FWHM resolution metrics, while DAS + UNET achieved the best CNR values. Our method achieved the second best CNR values, providing a good combination between resolution and contrast.

The publicly available PICMUS challenge [47] also allowed comparison of the proposed method to recently published beamforming methods evaluated on this dataset. The compared algorithms include methods based on the Coherence Factor (CF) [61] combination of MV and CF [62], and a method based on a Sparsity Prior (SP) [15]. All methods were tested on the same PICMUS datasets using the same resolution (axial and lateral FWHM) and contrast (CNR) metrics. As can be seen in Table 3 the proposed method achieved the best axial resolution results, while MV-CF [62] attained the best lateral resolution. In terms of contrast, our method achieved the best result. However, a better comparison would be to apply these methods with a following despeckling method such as DPAD, OBNMLF or ADMSS.

4.3. *In vivo data*

For additional qualitative inspection of the proposed method it was applied to an *in-vivo* carotid data set. Longitudinal views of the carotid obtained by the compared methods are presented in Fig. 13. One can observe that the DPAD, OBNLM and ADMSS methods reduce the speckle but it seems that they also smear some of the details. Our proposed intermediate output produces an image which suppresses most of the scatterers and amplifies the strong scatterers. However, for this dataset, the DAS + UNET and our networks' final output do not

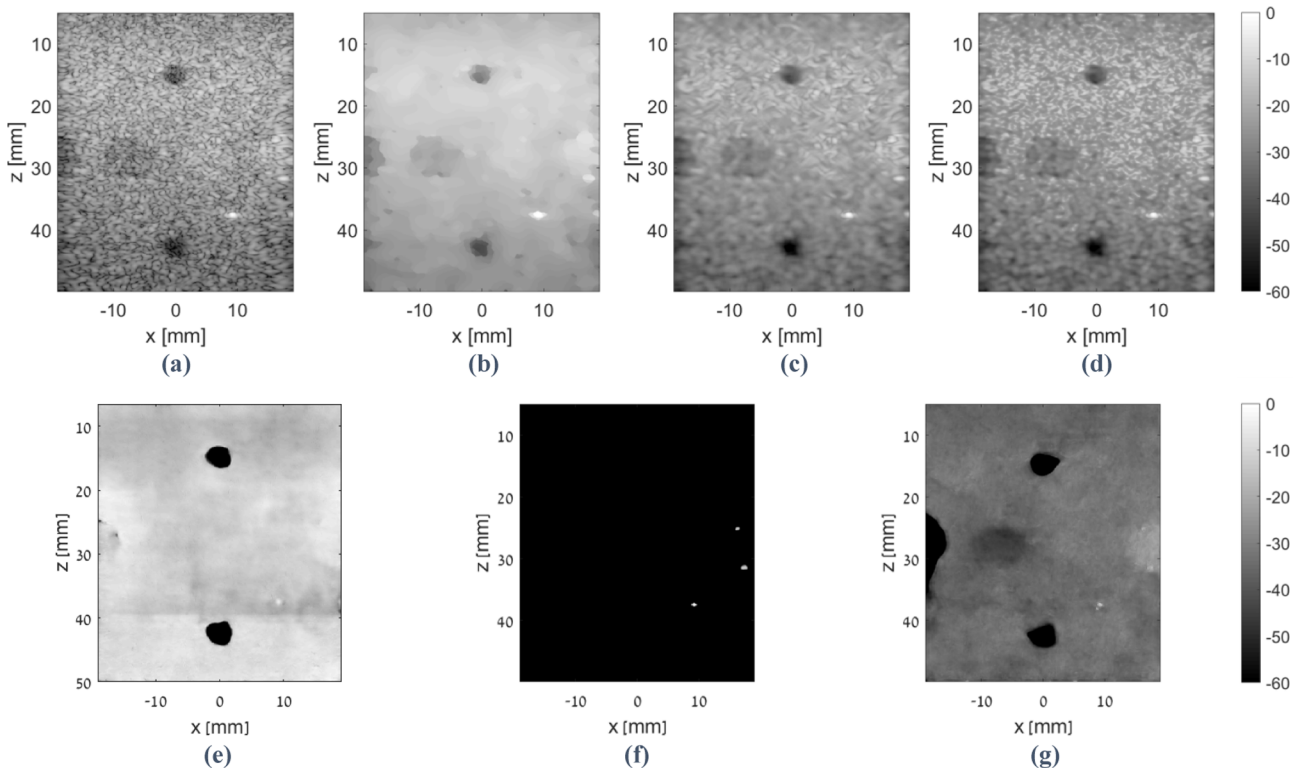

Fig. 12. Reconstruction results for the experimental contrast evaluation dataset. (a) DAS (b) DPAD (c) OBNMLF (d) ADMSS (e) DAS + UNET (f) Proposed network – intermediate output (g) Proposed network – final output.

Table 3
Comparison between earlier methods applied to the PICMUS dataset and the proposed method in terms of axial and lateral FWHM resolution (A = Axial, L = Lateral) and CNR for the different datasets (SR, SC, ER, EC).

Method	SR-FWHM		SC	ER-FWHM		EC
	A	L	CNR	A	L	CNR
DAS	0.4	0.82	9.96	0.57	0.89	8.15
CF [61]	0.23	0.5	11.93	0.31	0.59	9.65
MV-CF [62]	0.24	**0.08**	17.19	0.45	**0.2**	13.05
SP [15]	0.33	0.46	16.3	0.31	0.51	11.6
Proposed -Final Output	**0.19**	0.27	**22.63**	**0.22**	0.35	**16.75**

give better results than competitors, as they lose important details. While Fig. 13(g) do provide better segmentation of the artery itself (the black region in the image), the intima-media complex is not seen well, and it is difficult to estimate its thickness. This structure can be seen in the intermediate output (Fig. 13(f)), but it is not segregated in the final output.

5. Summary and future research directions

An end-to-end deep convolutional neural network was presented, preforming all the signal processing stages of ultrasonic PWI. The network is composed of two main parts. The first performs the beamforming stage based on a set of depth dependent learned spatial filters. The second is responsible for speckle reduction, performed by a UNET-like network. The network was trained in two stages. The first involved just the beamforming part of the network and was based on simulations of sparsely distributed scatterers. The full network was then trained using images containing clusters of fully resolved speckle intensities. Evaluation on simulations and experimental datasets obtained from the PICMUS challenge indicate that the proposed method is an attractive choice in terms of resolution and contrast, outperforming DAS followed by several common despeckling methods. However, the final output of the proposed method sometimes includes segmentation artifacts, especially in segmenting delicate lines like the intima-media complex in Fig. 13. One way to cope with this is to consider in diagnosis both the intermediate output, which is sensitive to isolated scatterers, and the final output, which is better in region segmenting in most cases. Future research avenues that we believe will reduce these artifacts include:

1. Training with a richer prior, including thin strips: Training in this work was done with simulated images containing ellipse-shaped areas, which accurately represent convex hull clusters, but not thin and delicate strips. Adding such strips to the simulated data is likely to help in establishing a more general prior, and help in segmenting regions like the intima-media complex.
2. Training with real ultrasonic scans: Network training here was based on images synthesized with Field-II software simulation [54] according to a pre-defined prior. Training with real ultrasonic scans will provide a better prior, reflecting the domain structure distribution. In addition such data will include phenomena not simulated by the Field-II software like ultrasonic shadowing, inter scatterer interactions, nonlinear effects, different kinds of noise, and velocity variations.
3. Multiple plain waves: This work dealt with a single plane wave sent in 0^o angle. Extending the method to multiple plane waves at different angles may lead to significant additional improvements.
4. Multiple spatial filters: The depth dependent spatial filtering layer was trained based on the assumption of constant wave velocity. Extending the network to include several filters at each depth may enable more robust results.

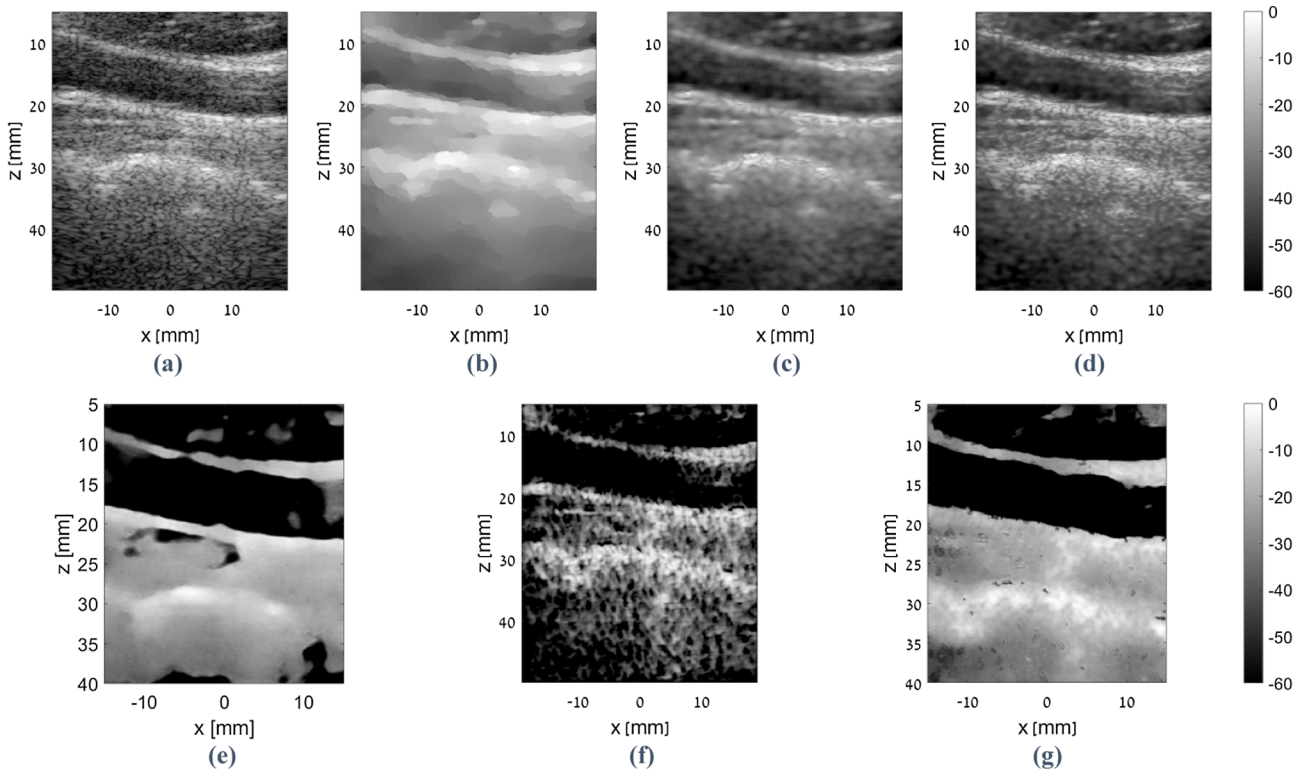

Fig. 13. In vivo carotid artery scan in longitudinal view. (a) DAS (b) DPAD (c) OBNMLF (d) ADMSS (e) DAS + UNET (f) Proposed network – intermediate output (g) Proposed network – final output.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the PAZI Foundation, Israel through Grant no. ID63-2018.

References

[1] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, M. Fink, Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 56 (3) (2009) 489–506.
[2] B. Denarie, T.A. Tangen, I.K. Ekroll, N. Rolim, H. Torp, T. Bjåstad, L. Lovstakken, Coherent plane wave compounding for very high frame rate ultrasonography of rapidly moving targets, IEEE Trans. Med. Imaging 2 (7) (2013) 1265–1276.
[3] B. Madore, P.J. White, K. Thomenius, G.T. Clement, Accelerated focused ultrasound imaging, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 56 (12) (2009).
[4] M. Karaman, P. Li, M. O'Donnell, Synthetic aperture imaging for small scale systems, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 42 (3) (1995) 429–442.
[5] J.A. Jensen, S.I. Nikolov, K.L. Gammelmark, M.H. Pedersen, Synthetic aperture ultrasound imaging, Ultrasonics 44 (2006) pp. e-e15.
[6] J. Capon, High-resolution frequency-wavenumber spectrum analysis, Proc. IEEE 57 (8) (1969) 1408–1418.
[7] C.C. Nilsen, S. Holm, Wiener beamforming and the coherence factor in ultrasound imaging, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 57 (6) (2010).
[8] H. Hasegawa, H. Kanai, Effect of subaperture beamforming on phase coherence imaging, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 61 (11) (2014) 1779–1790.
[9] J. Zhao, Y. Wang, J. Yu, W. Guo, T. Li, Y. Zheng, Subarray coherence based post-filter for eigenspace based minimum variance beamformer in ultrasound plane-wave imaging, Ultrasonics 65 (2016) 23–33.
[10] G. Matrone, A. Savoia, G. Caliano, G. Magenes, The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging, IEEE Trans. Med. Imaging 34 (4) (2015) 940–949.
[11] A.M. Moubark, Z. Alomari, S. Harput, D.M. Cowell, S. Freear, Enhancement of contrast and resolution of B-mode plane wave imaging (PWI) with non-linear filtered delay multiply and sum (FDMAS) beamforming, IEEE International Ultrasonics Symposium (IUS), 2016, pp. 1–4.
[12] J.A. Jensen, P. Gori, Spatial filters for focusing ultrasound images, IEEE International Ultrasonics Symposium (IUS), 2001, pp. 1507–1511.
[13] K. Kim, J. Liu, M.F. Insana, Efficient array beam forming by spatial filtering for ultrasound B-mode imaging, J. Acoust. Soc. America 120 (2) (2006) 852–861.
[14] Y. Chen, Y. Lou, J. Yen, Dynamic transmit-receive beamforming by spatial matched filtering for ultrasound imaging with plane wave transmission, Ultrasonic Imaging 39 (4) (2017) 207–223.
[15] E. Ozkan, V. Vishnevsky, O. Goksel, Inverse problem of ultrasound beamforming with sparsity constraints and regularization, IEEE Trans. Ultrasonics Ferroelect. Frequency Control (2017).
[16] T. Szasz, A. Basarab, D. Kouamé, Beamforming through regularized inverse problems in ultrasound medical imaging, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 63 (12) (2016) 2031–2044.
[17] P.M. Shankar, A general statistical model for ultrasonic backscattering from tissues, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 47 (3) (2000) 727–736.
[18] G.F. Pinton, J.J. Dahl, G.E. Trahey, Rapid tracking of small displacements with ultrasound, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 53 (6) (2006) 1103–1117.
[19] O.V. Michailovich, A. Tannenbaum, Despeckling of medical ultrasound images, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 53 (1) (2006) 64–78.
[20] Y. Yu, S.T. Acton, Speckle reducing anisotropic diffusion, IEEE Trans. Image Process 11 (11) (2002) 1260–1270.
[21] S. Aja-Fernández, C. Alberola-López, On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering, IEEE Trans. Image Process 15 (9) (2006) 2694–2701.
[22] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Martin-Fernandez, C. Alberola-López, S. Aja-Fernández, Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images, IEEE Trans. Image Process 24 (1) (2014) 345–358.
[23] Y. Yue, M.M. Croitoru, A. Bidani, J.B. Zwischenberger, J.W. Clark, Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images, IEEE Trans. Image Process 25 (3) (2006) 297–311.
[24] A. Khare, M. Khare, Y. Jeong, H. Kim, M. Jeon, Despeckling of medical ultrasound images using daubechies complex wavelet transform, Signal Process. 90 (2) (2010) 428–439.
[25] C. Kervrann, J. Boulanger, P. Coupé, Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal, International Conference on Scale Space and Variational Methods in Computer Vision, 2007, pp. 520–532.
[26] P. Coupé, P. Hellier, C. Kervrann, C. Barillot, Nonlocal means-based speckle filtering for ultrasound images, IEEE Trans. Image Process 18 (10) (2009) 2221–2229.
[27] L. Zhu, C. Fu, M.S. Brown, P. Heng, A non-local low-rank framework for ultrasound speckle reduction, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 5650–5658.
[28] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[29] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep

convolutional neural networks, Advances in neural information processing systems, 2012, pp. 1097–1105.
[30] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, Imagenet large scale visual recognition challenge, Int. J. Comput. Vision 115 (3) (2015) 211–252.
[31] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[32] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
[33] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015, pp. 234–241.
[34] H.C. Burger, C.J. Schuler, S. Harmeling, Image denoising: Can plain neural networks compete with BM3D? Proceedings of the IEEE conference on computer vision and pattern recognition, 2012, pp. 2392–2399.
[35] J. Xie, L. Xu, E. Chen, Image denoising and inpainting with deep neural networks, Advances in Neural Information Processing Systems, 2012, pp. 341–349.
[36] L. Xu, J.S. Ren, C. Liu, J. Jia, Deep convolutional neural network for image deconvolution, Advances in Neural Information Processing Systems, 2014, pp. 1790–1798.
[37] C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Machine Intelligence 38 (2) (2016) 295–307.
[38] J. Kim, J. Kwon Lee, K. Mu Lee, Accurate image super-resolution using very deep convolutional networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1646–1654.
[39] G. Wang, A perspective on deep imaging, IEEE access, 2016, pp. 8914–8924.
[40] D. Perdios, A. Besson, M. Arditi, J. Thiran, A deep learning approach to ultrasound image recovery, IEEE International Ultrasonics Symposium (IUS), 2017, pp. 1–4.
[41] A.C. Luchies, B.C. Byram, Deep neural networks for ultrasound beamforming, IEEE Trans. Med. Imaging 37 (9) (2018) 2010–2021.
[42] M. Gasse, F. Millioz, E. Roux, D. Garcia, H. Liebgott, D. Friboulet, High-quality plane wave compounding using convolutional neural networks, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 64 (10) (2017) 1637–1639.
[43] S. Vedula, O. Senouf, A.M. Bronstein, O.V. Michailovich and M. Zibulevsky, "Towards ct-quality ultrasound imaging using deep learning," arXiv preprint arXiv:1710.06304, 2017.
[44] D. Mishra, S. Chaudhury, M. Sarkar, A.S. Soin, Ultrasound image enhancement using structure oriented adversarial network, IEEE Signal Process Lett. 25 (9) (2018) 1349–1353.
[45] F. Dietrichson, E. Smistad, A. Ostvik, L. Lovstakken, "Ultrasound speckle reduction using generative adversial networks, IEEE International Ultrasonics Symposium (IUS) 2018, 2018, pp. 1–4.
[46] D. Hyun, L.L. Brickson, K.T. Looby, J.J. Dahl, Beamforming and speckle reduction using neural networks, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 66 (5) (2019) 898–910.
[47] H. Liebgott, A. Rodriguez-Molares, F. Cervenansky, J.A. Jensen, O. Bernard, "Plane-wave imaging challenge in medical ultrasound, IEEE International Ultrasonics Symposium (IUS), 2016, pp. 1–4.
[48] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," arXiv preprint arXiv:1603.07285, 2016.
[49] J.H. Chang, J.T. Yen, K.K. Shung, A novel envelope detector for high-frame rate, high-frequency ultrasound imaging, IEEE Trans. Ultrasonics Ferroelect. Frequency Control 54 (9) (2007) 1792.
[50] S.A. Tretter, Communication System Design Using DSP Algorithms: With Laboratory Experiments for the TMS320C6713TM DSK, Springer Science & Business Media, 2008.
[51] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.
[52] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, Proceedings of the 27th international conference on machine learning (ICML), 2010, pp. 807–814.
[53] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[54] J.A. Jensen, Field: A program for simulating ultrasound systems, 10th Nordicbaltic Conference on Biomedical Imaging, 1996, pp. 351–353 Supplement 1, Part 1.
[55] D.P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.
[56] F. Chollet, "Keras: Deep learning library for theano and tensorflow," URL: , vol. 7, no. 8, 2015. https://Keras.io/k.
[57] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean and M. Devin, "TensorFlow: Large-scale machine learning on heterogeneous systems. " URL, 2015. https://Www.tensorflow.org.
[58] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 249–256.
[59] J.A. Jensen, P. Munk, Computer phantoms for simulating ultrasound B-mode and CFM images, Acoustical imaging, Anonymous, Springer, 1997, pp. 75–80.
[60] O.M.H. Rindal, A. Austeng, Double adaptive plane-wave imaging, IEEE International Ultrasonics Symposium (IUS), 2016, pp. 1–4.
[61] C. Hu, M. Li, Improved plane wave imaging with retrospective transmit focusing and weighting based post filtering, IEEE International Ultrasonics Symposium (IUS), 2016, pp. 1–4.
[62] A.M. Deylami, J.A. Jensen, B.M. Asl, An improved minimum variance beamforming applied to plane-wave imaging in medical ultrasound, IEEE International Ultrasonics Symposium, 2016, pp. 1–4.