High-resolution image reconstruction for portable ultrasound imaging devices

Ruoyao Wang[1,2†], Zhenghan Fang[3†], Jiaqi Gu[4], Yi Guo[2,5], Shicong Zhou[6], Yuanyuan Wang[2,5*], Cai Chang[6] and Jinhua Yu[2,5*]

Abstract

Pursuing better imaging quality and miniaturizing imaging devices are two trends in the current development of ultrasound imaging. While the first one leads to more complex and expensive imaging equipment, poor image quality is a common problem of portable ultrasound imaging systems. In this paper, an image reconstruction method was proposed to break through the imaging quality limitation of portable devices by introducing generative adversarial network (GAN) model into the field of ultrasound image reconstruction. We combined two GAN generator models, the encoder-decoder model and the U-Net model to build a sparse skip connection U-Net (SSC U-Net) to tackle this problem. To produce more realistic output, stabilize the training procedure, and improve spatial resolution in the reconstructed ultrasound images, a new loss function which combines adversarial loss, L1 loss, and differential loss was proposed. Three datasets including 50 pairs of simulation, 40 pairs of phantom, and 72 pairs of in vivo images were used to evaluate the reconstruction performance. Experimental results show that our SSC U-Net is able to reconstruct ultrasound images with improved quality. Compared with U-Net, our SSC U-Net is able to preserve more details in the reconstructed images and improve full width at half maximum (FWHM) of point targets by 3.23%.

Keywords: Image reconstruction, Generative adversarial networks, Portable ultrasound imaging system

1 Introduction

Ultrasound has become an indispensable imaging technology in current clinical medicine. Compared with commonly used imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography/single-photon emission computed tomography (SPECT/PET), ultrasound imaging is real-time, portable, inexpensive, and free of ionizing radiation risks [1]. Over the past decades, ultrasound imaging has gone through rapid development and presented two trends. Since the major limitation of ultrasonography is the relatively low imaging quality, considerable effort has been made to improve the resolution and contrast, or reduce the artifacts of the ultrasound imaging [2–4]. Benefiting from major development of imaging techniques, the ultrasound imaging quality has obtained considerable improvement. Developed imaging algorithms or signal processing processes, however, naturally lead to more complex and expensive ultrasound equipment. On the other hand, in order to take full use of the portability of ultrasound imaging, the other trend is to minimize or simplify the imaging equipment for the wide range of applications such as family examination or health care in extreme environment [5–7]. Due to the size limit of the equipment, the imaging quality is further degraded in portable ultrasound imaging system.

Conventional ultrasound imaging systems usually weight hundreds of kilograms. Most of ultrasound imaging systems equipped with wheels, which allows a certain kind of mobility. For example, bedside ultrasound has been used in biopsy guidance [8], intraoperative navigation [9], and obstetrical monitoring [10]. The large size and heavy weight of conventional ultrasound equipment, however, prohibit the usage of ultrasound imaging in other out-of-hospital arenas. With recent advances in integrated circuit, it appeared lightweight pocket-size ultrasound imaging devices, which have been used in

* Correspondence: yywang@fudan.edu.cn; jhyu@fudan.edu.cn
Ruoyao Wang and Zhenghan Fang have the same contribution to this paper.
†Ruoyao Wang and Zhenghan Fang contributed equally to this work.
²Department of Electronic Engineering, Fudan University, Shanghai, China
Full list of author information is available at the end of the article

Wang *et al. EURASIP Journal on Advances in Signal Processing* (2019) 2019:56

Page 2 of 12

emergency aid on the accident scene, disaster relief, military operation, health care in spaceship, family doctor scenario, etc. [11, 12]. Despite the wide application possibilities, the poor imaging quality became a major limitation of the portable ultrasound imaging system. Therefore, it is of great interest to improve the imaging quality of portable ultrasound equipment.

There are three aspects concerning ultrasound imaging quality, namely, spatial resolution, contrast, and noise level. Compared with the traditional normal-size imaging devices, portable equipment typically produces images with lower spatial resolution, lower contrast, and greater noise. The poor imaging quality not only hinders the doctors from giving confident diagnosis but also misleads the doctors in making wrong decision or operation in emergency treatment. As a result, the poor imaging quality has become the major obstacle to the development and further application of portable ultrasound equipment. This motivates us to propose an ultrasound image reconstruction method to improve the imaging quality of portable equipment in terms of resolution, contrast, and noise reduction.

In the last three decades, many methods have been proposed for quality improvement in ultrasound imaging. Beamforming is a commonly used method to improve lateral/axial resolution or contrast of the imaging. By creating spatial selectivity of signals received from or sent to a transducer array, beamforming can produce imaging signals with narrow main-lobe, suppressed side-lobes, dynamic focus, and reduced speed of sound errors [13]. The representative beamforming algorithms include delay-and-sum (DAS), minimum variance (MV) [14, 15], and eigenspace-based MV (ESMV) [16, 17]. Adaptive beamforming such as MV or ESMV is able to provide resolution and contrast improvement around 30% or above than traditional DAS beamforming [18]. Besides beamforming methods, some deep learning methods are introduced recently to reconstruct images from radio frequency (RF) signals. In [19], Nair et al. designed a fully convolutional neural network to segment anechoic cysts directly from RF signals without beamforming. Luchies et al. [20] reconstructed ultrasound images from RF signals with a deep neural network and observed a better reconstruction quality compared with DAS beamforming. Although advanced beamforming methods and deep learning methods based on RF signals could improve the imaging quality successfully, this group of methods involve the complex calculation on RF signals which are hardly obtained in commercially available ultrasound imaging equipment.

Compared with the beamforming methods on RF signal domain, image reconstruction methods on image domain are more convenient and versatile. Yang et al. [21] used a variation of pixel compounding method to reconstruct a high-resolution image from a sequence of ultrasound images acquired with random motion. Taxt and Jirík [22] proposed a noise robust devolution method that deconvolved the first and second harmonic images separately, resulting in higher resolution images with reduced speckle noise. In [23], Chen et al. proposed a compressive deconvolution framework to reconstruct enhanced RF images by optimization method of the alternating direction method of multipliers. The main assumption that has been used in [23] is that point spread function (PSF) of the ultrasound imaging is spatially invariant and can be estimated from RF signal. As a result, for reconstructing a high-quality image, image reconstruction–based methods usually need more information besides a low-quality image. A few measurements with random motion, fundamental and harmonic images, and even parameter estimation from RF signal are required in [21–23], respectively.

Speckle noise reduction is also an important aspect of image quality improvement since the existence of speckle considerably lowers the image contrast and blurs image details. Many speckle reduction techniques have been proposed such as frequency/spatial compounding [24, 25], spatial filtering [26, 27], and multiscale analysis [28, 29]. Although some methods are able to reduce speckle noise effectively and helpful to image analysis task such as image segmentation, registration, and object detection, this group of methods always tries to strike a balance between noise reduction and detail preservation. Furthermore, speckle noise reduction has no contribution to the resolution which is the most important index of imaging quality.

Abovementioned methods such as beamforming, image reconstruction, and noise reduction tend to focus on one or two aspects of image quality. In this paper, we tried to improve image quality in all aspects by using deep learning method to generate high-quality images. Previous works proved that deep learning methods can be applied to medical image generation [30–32] and usually outperform traditional methods. Specifically for ultrasound image reconstruction, comparing with normal-size ultrasound imaging equipment, the imaging quality of portable equipment is degraded in resolution, contrast, and heavy noise jointly. To reconstruct a high-resolution ultrasound image from a low-resolution one is similar to an image-to-image translation task [33]. We followed some previous works dealt with similar problems using GAN (Generative adversarial networks) in this paper. An image reconstruction method based on GAN was proposed to break through the imaging quality limitation of portable ultrasound devices. For the task at hand, the GAN-based method has the following advantages: (1) Multi-level nonlinear mapping

Fig. 1 Flow chart of the reconstruction algorithm
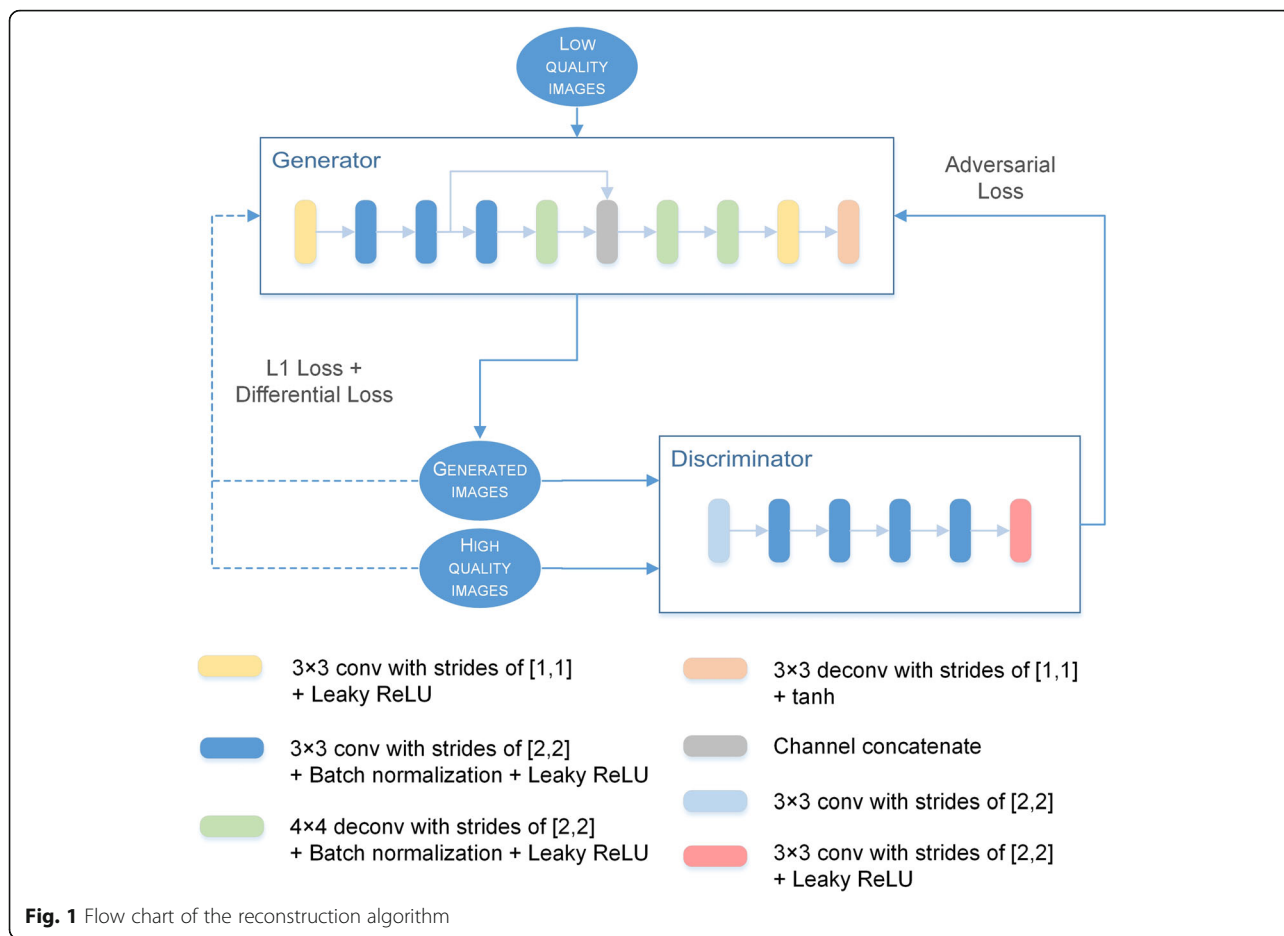
relationship between the low-quality images and the high-quality images can be extracted by learning model, which therefore has potential to improve imaging quality in multiple aspects. (2) The feature extractors are automatically learned from actual ultrasound images, not human designed, and therefore more representative and adaptive to data. (3) The discriminator used in GAN is able to improve imaging quality visually. (4) Once the model is trained, the reconstruction procedure is a one-step and feedforward process, which is therefore more direct and efficient than some other methods that involve iterated calculations, also more suitable for real-time ultrasound image processing. (5) The fast developing technology on hardware implementation of neural networks allows our method to be implemented on small and portable hardware like FPGAs, and thus easily incorporated into the current portable ultrasound equipment.

The rest of the paper is organized as follows: Methods describes the proposed method and the experimental data, Experiments shows the experimental results, and Results and discussion concludes our work.

2 Methods

2.1 Network architecture

In this study, we proposed a GAN model to reconstruct high-resolution images for portable ultrasound imaging devices. The network architectures of the GAN models used in this study are shown in Fig. 1.

2.1.1 Generator with sparse skip connections

There are many choices to build the generator of a GAN. One choice is to use an encoder-decoder model as the GAN generator [34]. In an encoder-decoder model, an encoder is defined as a 3×3 convolution layer followed by an activation layer and a batch-normalization layer. The stride of the convolution is [2, 2] in order to downsample the image. The notation $[x, y]$ means that for a two dimension convolution, the stride is of the convolution kernel is x and y for the first and second dimension respectively. A decoder has a similar structure to the encoder, but uses a 4×4 deconvolution instead in order to avoid checkerboard artifacts [35]. The stride of the deconvolution is [2, 2] in order to upsample the image. An encoder-decoder model usually uses the same number of encoders and decoders. An
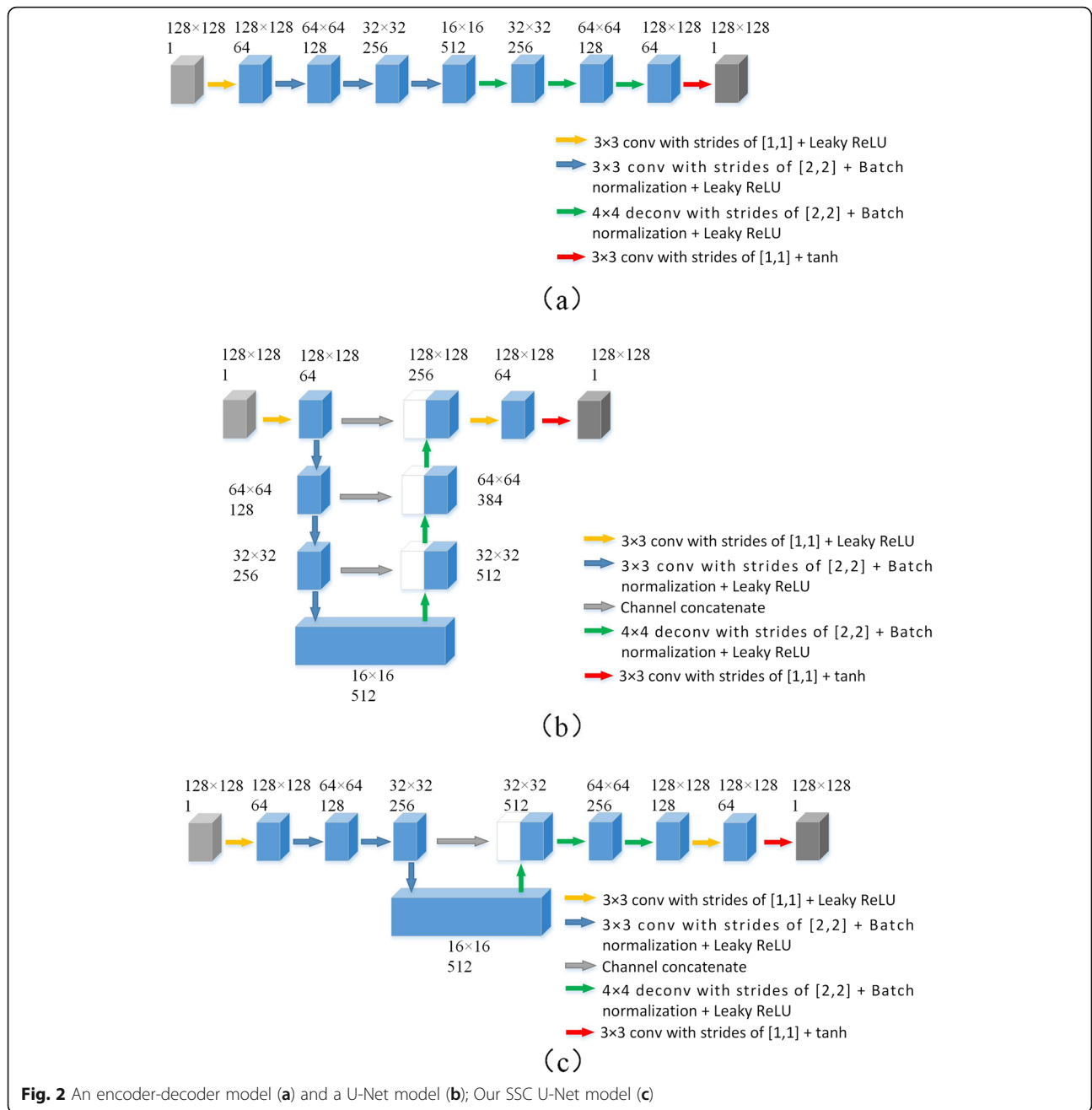
Fig. 2 An encoder-decoder model (**a**) and a U-Net model (**b**); Our SSC U-Net model (**c**)

input image is first downsampled by encoders, and then upsampled to restore the original size by decoders. The encoder-decoder model we used in this paper is given in Fig. 2(a). It has three encoders and decoders. The input of the model is a 128 × 128 one channel patch. The activation function we used is leaky ReLU.

However, as is demonstrated in [33], a bottleneck exists in the structure of the encoder-decoder model, which limits the sharing of low-level information between input and output. In [33], the authors proposed a U-Net model to allow more low-level information to

pass from the input to the output. To do this, the U-Net model adds skip connections between mirrored encoders and decoders compared with an encoder-decoder model. The U-Net model we used is given in Fig. 2(b). The U-Net model has been successfully applied to the super-resolution reconstruction of many medical images, such as MRI [30, 32], plane wave ultrasound images [36], and CT [37]. However, as low-quality ultrasound images have many speckles and artifacts, applying a U-Net model to the ultrasound image super-resolution reconstruction task raises a new issue: sharing all low-level

information between the input and the output will bring speckles and artifacts in low-resolution images into reconstructed high-resolution images. This is because there is a shallow feedforward path through the top skip connections which extract few features from the input images.

In order to maintain the structure information in the low-resolution image, and meanwhile not bring speckles and imaging artifacts in the low-resolution image to the high-resolution one, we designed a new generator which only concatenated the output of the third encoder to the input of the third decoder. This design keeps the benefits of the U-Net while reduces the low-level information parameters. We call our model a sparse skip connection U-Net (SSC U-Net). Our network is shown in Fig. 2(c).

2.1.2 Discriminator and training strategy

The discriminator uses local patches of the origin image during the training process. This strategy is based on the assumption that pixels from different patches are independent. It encourages the discriminator to model high-frequency details [38]. Other benefits of local patching are that it helps enlarge the data set and save memory resources during the training process. This idea is commonly accepted in tasks like image style transference [39]. The patch size in our network is 128 × 128.

Our training strategy follows the approach in [40]. We train the generator first. The discriminator is trained then with the real images and the images generated by the generator. We use Adam solver in our method.

2.2 Objective

A cross entropy loss is usually used in GAN training. We call this loss adversarial loss:

$$L_{GAN}(G,D) = E_{x,y \sim P_{data}(x,y)}\big[\ \mathrm{log}D(x,y)\big] + \\ E_{x \sim P_{data}(x)}\big[\ \mathrm{log}(1{-}D(G(x)))\big] \quad (1)$$

where x refers to the input vector and y refers to the output vector. D refers to the discriminator and G refers to the generator. L refers to the loss function and E refers to expectation. The adversarial loss is used to let the generator generates images as close to the real images as possible.

In our task, we needed to reconstruct a high-resolution image according to the input low-resolution images. This is a supervised learning, so we implemented L1 loss in our model to maintain pixel-wise similarity [41]. Previous work [42] confirms that it is possible to mix traditional loss with adversarial loss in GAN training. L1 loss helps stabilize the training and preserves low-frequency information from images. L1 loss defined as follows:
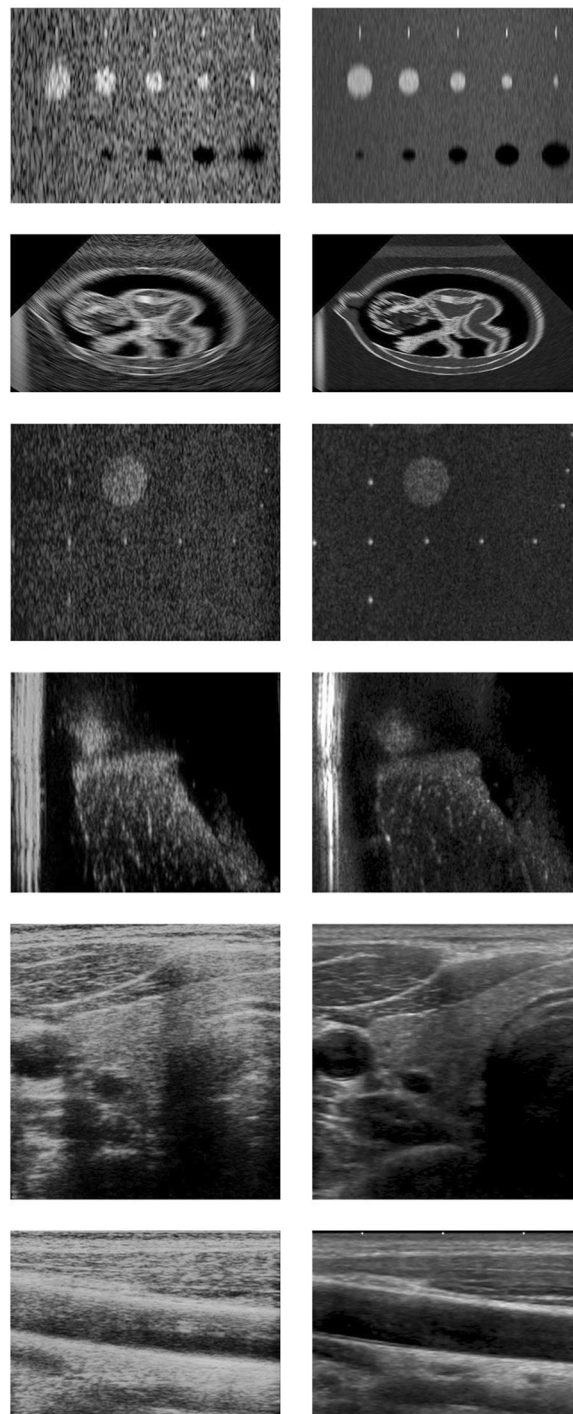

Fig. 3 Examples of experimental data

$$L_{L1}(G) = E_{x,y \sim P_{data}(x,y)}\big[\|y{-}G(x)\|_1\big] \quad (2)$$

However, L1 loss will result in blurring in output [41]. Texture and speckle noises, for example, are likely to be blurred due to L1 loss. While L1 loss is essential to keep low-frequency information, we introduced a differential

Table 1 Results of the three models on simulation dataset

	PSNR	SSIM	MI	FWHM	CR
	Mean (Std.)	Mean (Std.)	Mean (Std.)	Mean (Std.)	Mean (Std.)
Low quality	17.74 (0.63)	0.22 (0.07)	0.36 (0.25)	22.6 (1.64)	0.67 (0.01)
Encoder-decoder	24.02 (1.79)	0.50 (0.11)	0.38 (0.08)	17.6 (1.75)	0.64 (0.07)
U-Net	25.69 (1.38)	0.68 (0.04)	0.55 (0.10)	15.5 (2.41)	0.64 (0.05)
SSC U-Net	26.15 (0.71)	0.64 (0.01)	0.60 (0.01)	15.0 (1.36)	0.67 (0.02)

loss to preserve the sharpness of edges in the generated images. The differential loss is defined as the absolute difference between the sum of the vertical and horizontal gradient of the generated image and the real image. The differential loss is given as follows:

$$L_{diff}(G) = E_{x,y \sim P_{data}(x,y)} \left[\sum_{i=1,2} | \frac{\partial y}{\partial x_i} - \frac{\partial}{\partial x_i} G(x) | \right] \quad (3)$$

Our final objective is:

$$G^* = \arg \min_G \max_D L_{GAN}(G,D) + \alpha L_{L1}(G) + \beta L_{diff}(G) \quad (4)$$

In our experiments, we choose $\alpha = 100$ and $\beta = 80$ in Eq. (4).

3 Experiments
3.1 Training datasets
Three datasets including 50 pairs of simulation, 40 pairs of phantom, and 72 pairs of in vivo images were used to train and test the GAN model. High-quality and low-quality images of phantom and in vivo data are generated from different devices and hence need to be registered. We align the images using a non-rigid image registration algorithm introduced in [43]. Mutual information is used as the similarity metric [30].

A total of 50 pairs of simulation data are generated by Field II ultrasound simulation program [44, 45]. All simulated data was generated by plane wave transmission. Two simulation models including cysts and fetus are used. The images of cyst phantoms are simulated with the following setting: number of transducer elements = 64, number of scanning lines = 50, number of scatterers = 100000 and central frequency = 3.5 MHz, 5 MHz, and 8 MHz. We simulated 20 images for each central frequency. The images of 3.5 and 5 MHz central frequency are used as low-quality images. The images of 8 MHz central frequency are averaged to get a high-quality image. For the fetus phantom, we simulate 10 low-quality images with central frequency = 5 MHz, number of transducer elements = 64, number of scanning lines = 128, and number of scatterers = 200000, and 10 high-quality images with central frequency = 8 MHz, number of transducer elements = 128, number of scanning lines = 128, and number of scatterers = 500000. According to [44, 45], these settings are able to simulate fully developed speckles. In all, we obtain 50 pairs of simulated images with 40 pairs of cysts and 10 pairs of fetus.

A total of 40 pairs of phantom data are generated from Vantage 64^{TM} research ultrasound system (Verasonics, Inc., USA) and mSonics MU1 (Youtu Technology, China). The phantoms include two CIRS
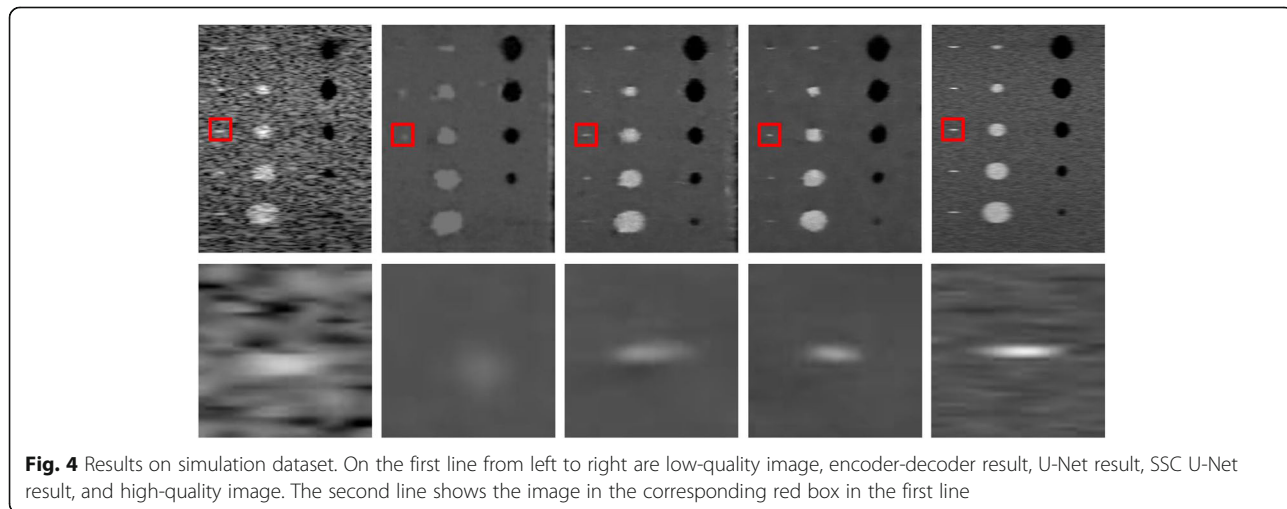
Fig. 4 Results on simulation dataset. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image. The second line shows the image in the corresponding red box in the first line
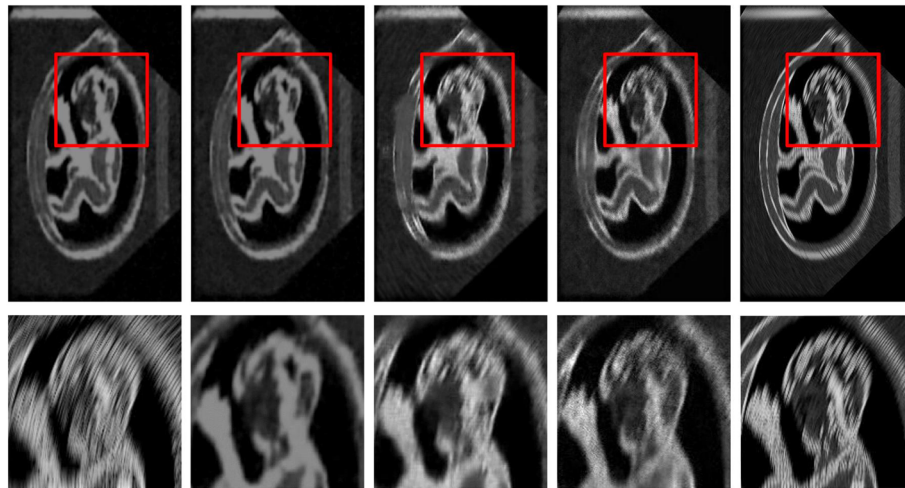
Fig. 5 Result examples of simulated fetus images. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image correspondingly. The second line shows the image in the corresponding red box in the first line

phantoms (Computerized Imaging Reference Systems, Inc., USA): ultrasound resolution phantom (model 044) and multi-purpose multi-tissue ultrasound phantom (model 040GSE), and two self-made pork phantoms. Verasonics Vantage 64 programmable ultrasound system is used for high-quality images and Msonics MU1 handheld ultrasound scanner for low-quality ones. The setting of Verasonics ultrasound system is as follows: central frequency = 7 MHz, dynamic range = 50 dB with multi-angle plane wave compounding method (20 angles from – 16 to 16°) and a 40-mm wide L11-4v transducer with 128 elements. The setting of Msonics MU1 handheld ultrasound scanner is as follows: central frequency = 7 MHz, gain = 70 dB, using a 40-mm wide L10-5 transducer with 128 elements. In all, we acquire 40 pairs of phantom images, 25 for CIRS phantoms, and 15 for pork phantoms.

Seventy-two pairs of in vivo data are generated from Toshiba Aplio 500 (Toshiba Medical Systems Corporation, Japan) and mSonics MU1 (Youtu Technology, China). Ultrasound images of thyroid from 50 subjects and images of carotid from 22 subjects are scanned. The clinical machine of Toshiba Aplio 500 with central frequency = 7.5 MHz and gain = 76 dB is used to generate the high-quality images. The parameter for portable machine Msonics MU1 is set as follows: central frequency = 6 MHz, gain = 95 dB, using a 40-mm wide L10-5 transducer with 128 elements. The focal depth of the ultrasound image acquisition is around 1 to 2 cm for in vivo data.

Examples of experimental data are shown in Fig. 3. The left and right column in Fig. 3 shows the low-quality and high-quality images respectively. From top to bottom are examples of simulation, phantom, and in vivo data respectively, two examples for each category.

3.2 Performance metrics

To evaluate the image reconstruction performance, we calculate peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and mutual information (MI) for all three datasets. Full width at half maximum (FWHM) of point targets and contrast resolution (CR) are calculated for simulated point object images and cyst images respectively.

PSNR measures the similarity of two images. If one is the ground-truth image, it measures the quality of the other one:

Table 2 Results of the three models on phantom dataset

	PSNR	SSIM	MI
	Mean (Std.)	Mean (Std.)	Mean (Std.)
Low quality	20.42 (1.54)	0.34 (0.05)	0.25 (0.22)
Encoder-decoder	24.42 (1.76)	0.50 (0.08)	0.37 (0.07)
U-Net	24.49 (2.39)	0.61 (0.09)	0.43 (0.11)
SSC U-Net	25.64 (1.79)	0.60 (0.13)	0.44 (0.10)

Table 3 Results of the three models on in vivo dataset

	PSNR	SSIM	MI
	Mean (Std.)	Mean (Std.)	Mean (Std.)
Low quality	12.45 (1.51)	0.22 (0.04)	0.45 (0.14)
Encoder-decoder	17.24 (0.72)	0.36 (0.03)	0.48 (0.11)
U-Net	18.37 (0.50)	0.38 (0.03)	0.50 (0.03)
SSC U-Net	18.52 (0.59)	0.38 (0.04)	0.50 (0.11)
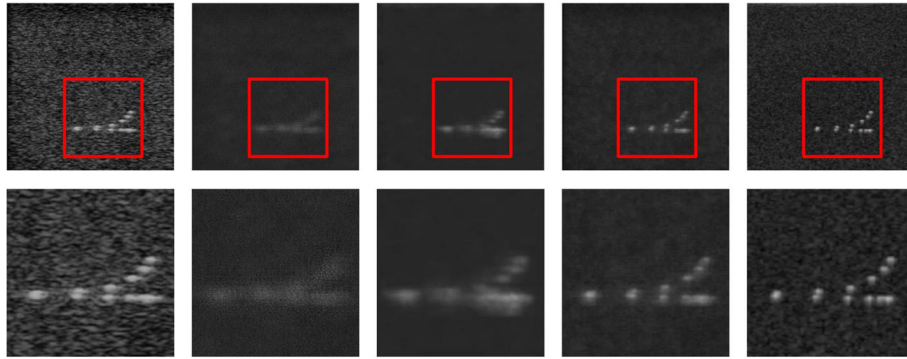
Fig. 6 Result examples of cyst phantom images. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image correspondingly. The second line shows the image in the corresponding red box in the first line

$$PSNR = 10 \times \log_{10}\left(\frac{(2^n-1)^2}{MSE}\right) \qquad (5)$$

where MSE is the mean squared error between reconstructed and ground-truth high-quality images, and for uint8 images, $n = 8$. The higher PSNR indicates the higher intensity similarity between reconstruction image and high-quality image.

SSIM measures the similarity of two images. It is defined as follows:

$$SSIM = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a{}^2 + \mu_b{}^2 + c_1)(\sigma_a{}^2 + \sigma_b{}^2 + c_2)} \qquad (6)$$

where a and b are windows on the reconstructed and ground-truth high-quality images respectively, μ_a and μ_b are average values of a and b, $\sigma_a{}^2$ and $\sigma_b{}^2$ are variances of a and b, and σ_{ab} is the covariance between a and b. c_1 and c_2 prevent the denominator from being zero. $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ where $k_1 = 0.01$ and $k_2 = 0.03$. The windows are isotropic Gaussian function with standard deviation of 1.5. The higher MSSIM represents the

higher structure similarity between reconstruction image and high-quality image.

MI indicates the mutual dependence between two images. It is defined as follows for a uint8 image:

$$MI(I, J) = \sum_0^{255} P_{IJ}(m) \log \frac{P_{IJ}(m)}{P_I P_J} \qquad (7)$$

where P_I and P_J represent the marginal probability distributions of I and J and P_{IJ} is the joint probability distribution of I and J.

CR is a measure of ability to distinguish difference in intensity in an image:

$$CR = \frac{|S_A - S_B|}{S_A + S_B} \qquad (8)$$

where S_A and S_B represent the average intensity values in two regions of interests (ROIs) with the same size. The higher CR represents the higher imaging contrast.

Since FWHM, CR is only applicable for simulation data, and PSNR, SSIM, and MI are used to evaluate the performance of algorithms in phantom data and in vivo data. The following tables in this part record
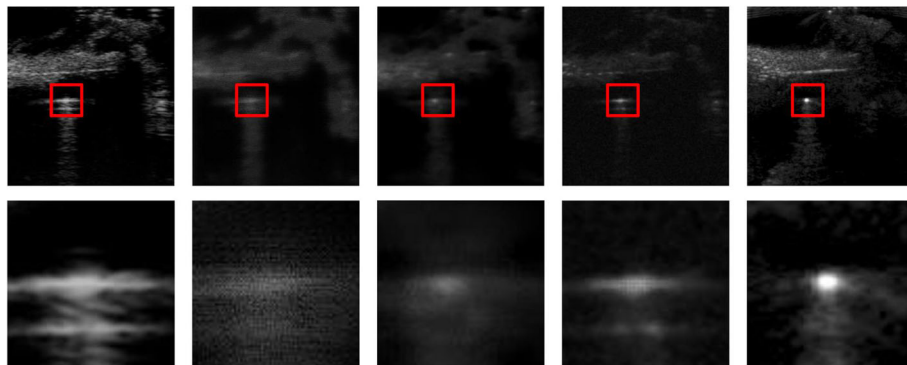
Fig. 7 Result examples of pork phantom images. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image correspondingly. The second line shows the image in the corresponding red box in the first line
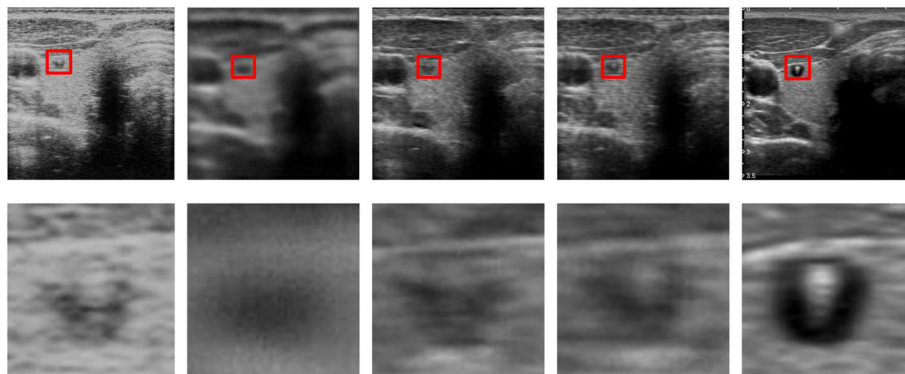
Fig. 8 Results of in vivo thyroid data. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image correspondingly. The second line shows the image in the corresponding red box in the first line

the results of 5-fold cross-validation experiments. Each dataset is randomly divided into five groups. One group is used as test data while other groups are used as train data each time.

3.3 Implementation
We implement our reconstruction model by Tensorflow. We use a Titan Xp Graphic Processing Unit (GPU) for training. We generated a 128×128 patch every ten pixels from an original image. The total number of data patches is 68,450 pairs of simulation patches, 38,480 pairs of phantom patches, and 73,736 pairs of in vivo patches. The input images are uint8 images. We rescaled their pixel intensity values to [− 1, 1]. The mean value of the training dataset is minused from the input data. Output images are added by the mean value of training dataset. Intensity values of the output images are clipped to [− 1, 1]. The images are then transformed into uint8 grayscale images. Learning rate of training is set to 0.00005 and follows an exponential decay to 0.000005. β_1 for Adam solver is set to 0.9.

4 Results and discussion
In this section, we report the results from the three models we tested: the encoder-decoder model, the U-Net model, and our SSC U-Net on simulation, phantom, and in vivo data.

4.1 Simulation data results
The simulation results are shown in Table 1. The SSC U-Net model out-performed the encoder-decoder model and the U-Net model with higher PSNR, MI, and CR and smaller FWHM. Though the U-Net model had the highest SSIM, we can observe that the generated images of the U-Net model were over-smoothed. This means that some high-frequency information was lost. The errors of the generated images and ground-truth images were averaged by over-smoothing and thus achieved a better SSIM score. Figure 4 gives examples of simulated cysts images. The point object in the red box reconstructed by SSC U-Net had a better resolution than other methods. Figure 5 gives examples of simulated fetus images. Details are clearer in the result generated by SSC U-Net.
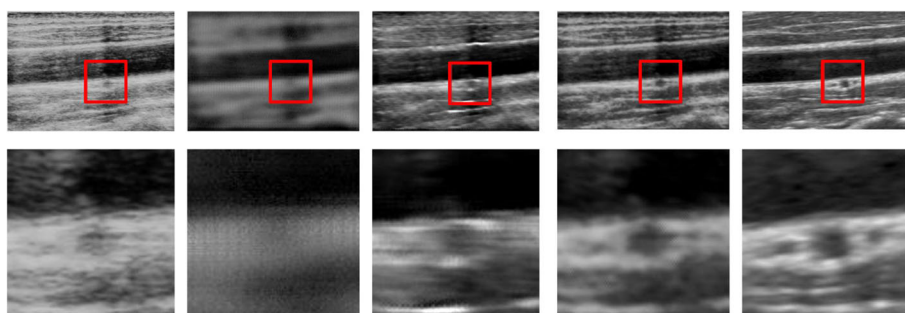
Fig. 9 Results of in vivo carotid data. On the first line from left to right are low-quality image, encoder-decoder result, U-Net result, SSC U-Net result, and high-quality image correspondingly. The second line shows the image in the corresponding red box in the first line
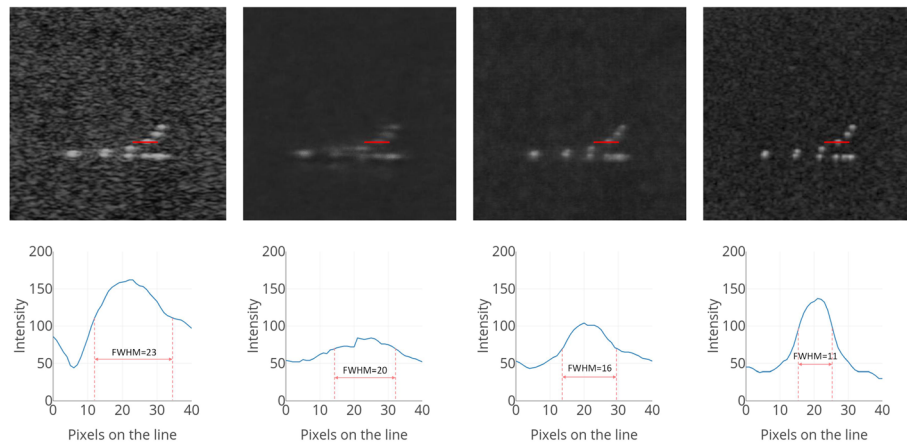
Fig. 10 The left column from top to bottom are a low-quality image, an image generated without differential loss, an image generated with differential loss, and a high-quality image. The right column shows the intensity on the red line of the corresponding image. Intensity values range from 0 to 255

4.2 Phantom data results

Table 2 shows the average performance of the three models on the phantom dataset. Our model maintained the highest PSNR and MI. The SSIM of the U-Net model is 1.6% more than our method. However, the U-Net model kept its drawbacks of over-smoothing. As is mentioned in section 4.1, over-smoothing may lead to a higher SSIM. Figure 6 gives an example of the CIRS phantoms results. The SSC U-Net model reconstructed the point objects well while other methods blurred these point objects. Figure 7 gives an example of pork phantoms results. The point object in the red box reconstructed by SSC U-Net had the highest resolution.

4.3 In vivo data results

Table 3 presents the results tested on the in vivo dataset. Our SSC U-Net achieved the best performance on PSNR, SSIM, and MI. Figures 8 and 9 give two examples of the generated images. Encoder-decoder models failed to generate a satisfying result on this dataset. Images generated by the U-Net smoothed some important details. For example, the calcification point in Fig. 8 and the cyst in Fig. 9 were blurred by the U-Net. In comparison, the result of our SSC-Net was better visually and kept more details.

4.4 Differential loss

The differential loss was introduced in our network to maintain sharpness of edges. The impact of the differential loss was investigated in an ablation experiment. We chose a horizontal line that crossed the center of a point object and plotted the intensity of pixels on the line in Fig. 10. We conclude from the images that with the help of differential loss, the reconstructed images achieve a higher resolution.

5 Conclusion

Image quality is vital for portable ultrasound imaging devices. In this paper, we proposed a new generative model called SSC U-Net to improve the image quality of portable ultrasound imaging devices. We tested our model on three datasets: simulation data, phantom data, and in vivo data. We compared our results with two other widely used GAN models: the encoder-decoder model and the U-Net model. Our experiment results show that our SSC U-Net model out-performed two other models in general on all three datasets. Images generated by our SSC U-Net had a better resolution and preserved more details than the other two methods.

Availability of data and materials
Please contact author for data requests.

Wang *et al. EURASIP Journal on Advances in Signal Processing*   (2019) 2019:56

Page 11 of 12

Author details
[1]University of Michigan, Ann Arbor, USA. [2]Department of Electronic Engineering, Fudan University, Shanghai, China. [3]University of North Carolina, Chapel Hill, USA. [4]University of Texas, Austin, USA. [5]Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai, China. [6]Department of Ultrasound, Fudan University Shanghai Cancer Center, Shanghai, China.

References
1. V. Chan, A. Perlas, *Atlas of Ultrasound-Guided Procedures in Interventional Pain Management*, 1st edn. (Springer, New York, 2011), pp. 13–19
2. I.K. Holfort, F. Gran, J.A. Jensen, Broadband minimum variance beamforming for ultrasound imaging. *IEEE Trans. Ultrason Ferroelectr Freq. Control* **56**(2), 314–325 (2009)
3. N. Deshpande, A. Needles, J.K. Willmann, Molecular ultrasound imaging: current status and future directions. *Clin Radiol* **65**(7), 567–581 (2010)
4. A. Perperidis, Postprocessing approaches for the improvement of cardiac ultrasound B-mode images: a review. *IEEE Trans Ultrason Ferroelectr Freq Control* **63**(3), 470–485 (2016)
5. P.B. McBeth, T. Hamilton, A.W. Kirkpatrick, Cost-effective remote iPhone-teathered telementored trauma telesonography. *J Trauma* **69**(6), 1597–1599 (2010)
6. A. Evangelista, V. Galuppo, J. Méndez, L. Evangelista, L. Arpal, C. Rubio, M. Vergara, M. Liceran, F. López, C. Sales, et al., Hand-held cardiac ultrasound screening performed by family doctors with remote expert support interpretation. *Heart* **102**(5), 376–382 (2016)
7. P Bornemann, G Bornemann, Military family physicians' perceptions of a pocket point-of-care ultrasound device in clinical practice, *Mil Med*, 179(12), 1474-1477(2014).
8. L. Fuso, F. Varone, D. Magnini, F. Baldi, G. Rindi, G. Pagliari, R. Inchingolo, A. Smargiassi, Role of ultrasound-guided transbronchial biopsy in the diagnosis of peripheral pulmonary lesions. *Lung Cancer* **81**(1), 60–64 (2013)
9. S. Satou, T. Aoki, J. Kaneko, Y. Sakamoto, K. Hasegawa, Y. Sugawara, O. Arai, T.M. MEng, K. Miura, N. Kokudo, Initial experience of intraoperative three-dimensional navigation for liver resection using real-time virtual sonography. *Surgery* **155**(2), 255–262 (2014)
10. S.V.D. Wal, S.J. Robson, S. Choong, Is bedside ultrasound now a routine part of specialist obstetrics and gynaecology outpatient consultation? *Australas J Ultrasound Med* **16**(4), 190–192 (2013)
11. B.P. Elson, E.R. Melnick, J. Li, Portable ultrasound for remote environments, Part I: Feasibility of field deployment. *J Emerg Med* **40**(2), 190–197 (2011)
12. B.P. Elson, E.R. Melnick, J. Li, Portable ultrasound for remote environments, part II: current indications. *J Emerg Med* **40**(3), 313–321 (2011)
13. S.H.C. Ortiz, T. Chiu, M.D. Fox, Ultrasound image enhancement: a review. *Biomed Signal Proces* **7**(3), 419–428 (2012)
14. J Capon, "High resolution frequency-wavenumber spectrum analysis," in *Proc. IEEE*, vol. 57, no. 8, pp. 1408-1418, Aug. 1969.
15. J.F. Synnevag, A. Austeng, S. Holm, Benefits of minimum-variance beamforming in medical ultrasound imaging. *IEEE Trans Ultrason Ferroelectr Freq Control* **56**(9), 1868–1879 (2009)
16. S. Mehdizadeh, A. Austeng, T.F. Johansen, S. Holm, Eigenspace based minimum variance beamforming applied to ultrasound imaging of acoustically hard tissues. *IEEE Trans Med Imag* **31**(10), 1912–1921 (2012)
17. X. Zeng, Y. Wang, J. Yu, Y. Guo, Beam-domain eigenspace-based minimum variance beamformer for medical ultrasound imaging. *IEEE Trans. Ultrason Ferroelectr Freq Control* **60**(12), 2670–2676 (2013)
18. B. Madore, F.C. Meral, Reconstruction algorithm for improved ultrasound image quality, *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*. **59**(2), 217–230 (2012)
19. A.A. Nair, T.D. Tran, A. Reiter, M.A.L. Bell, in *A deep learning based alternative to beamforming ultrasound images. IEEE International Conference on Acoustics, Speech and Signal Processing* (Calgary, IEEE, 2018), pp. 3359–3363
20. A.C. Luchies, B.C. Byram, Deep neural networks for ultrasound beamforming. *IEEE Trans. Med. Imag* **37**(9), 2010–2021 (2018)
21. Z. Yang, T.A. Tuthill, D.L. Raunig, M.D. Fox, M. Analoui, Pixel compounding: resolution-enhanced ultrasound imaging for quantitative analysis. *Ultrasound Med Biol* **33**(8), 1309–1319 (2007)
22. T. Taxt, R. Jirík, Superresolution of ultrasound images using the first and second harmonic signal. *IEEE Trans Ultrason Ferroelectr Freq Control* **51**(2), 163–175 (2004)
23. Z. Chen, A. Basarab, D. Kouamé, Compressive deconvolution in medical ultrasound imaging. *IEEE Trans Med Imag* **35**(3), 728–737 (2016)
24. J.R. Sanchez, M.L. Oelze, An ultrasonic imaging speckle-suppression and contrast-enhancement technique by means of frequency compounding and coded excitation. *IEEE Trans Ultrason Ferroelectr Freq Control* **56**(7), 1327–1339 (2009)
25. J. Opretzka, M. Vogt, H. Ermert, A high-frequency ultrasound imaging system combining limited-angle spatial compounding and model-based synthetic aperture focusing. *IEEE Trans Ultrason Ferroelectr Freq Control* **58**(7), 1355–1365 (2011)
26. Y. Yu, S.T. Acton, Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process* **11**(11), 1260–1270 (2002)
27. K. Krissian, C.F. Westin, R. Kikinis, K.G. Vosburgh, Oriented speckle reducing anisotropic diffusion. *IEEE Trans Image Process* **16**(5), 1412–1424 (2007)
28. 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 Med Imag* **25**(3), 297–311 (2006)
29. J. Kang, J.Y. Lee, Y. Yoo, A new feature-enhanced speckle reduction method based on multiscale analysis for ultrasound B-mode imaging. *IEEE Trans. Biomed Eng* **63**(6), 1178–1191 (2016)
30. D. Nie, R. Trulo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans Biomed Eng* **65**(12), 2720–2730 (2018)
31. G. Yang et al., DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. *IEEE Trans Med Imag* **37**(6), 1310–1321 (2018)
32. K. Bahrami, F. Shi, X. Zong, H.W. Shin, H. An, D. Shen, Reconstruction of 7T-Like Images From 3T MRI. *IEEE Trans Med Imag* **35**(9), 2085–2097 (2016)
33. O. Ronneberger, P. Fischer, T. Brox, in *U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer Assisted Intervention* (MICCAI, Cham, 2015), pp. 234–241
34. G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks. *Science* **313**(5786), 504–507 (2006)
35. A. Odena, V. Dumoulin, C. Olah, Deconvolution and checkerboard artifacts. *Distill* **1**(10), e3 (2016)
36. Z. Zhou, Y. Wang, J. Yu, Y. Guo, W. Guo, Y. Qi, High spatial–temporal resolution reconstruction of plane-wave ultrasound images with a multichannel multiscale convolutional neural network. *IEEE Trans Ultrason Ferroelectr Freq Control* **65**(11), 1983–1996 (2018)
37. C. You et al., CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE), *IEEE Trans. Med. Imag.*, 2019). https://doi.org/10.1109/TMI.2019.2922960
38. L.A. Gatys, A.S. Ecker, M. Bethge, in *Image style transfer using convolutional neural networks. IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, IEEE, 2016), pp. 2414–2423
39. D. Pathak, P. Krähenbühl, J. Donahue, T. Darrell, A.A. Efros, in *Context encoders: feature learning by inpainting. IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Las Vegas, 2016), pp. 2536–2544
40. I Goodfellow, J Pougetabadie, M Mirza, B Xu, D Warde-Farley, S Ozair,A Courville, Y Bengio, in *Conference and Workshop on Neural Information Processing Systems*, Generative Adversarial Nets(NIPS, Montreal, 2014), pp. 2672–2680.
41. P. Isola, J.Y. Zhu, T. Zhou, A.A. Efros, in *Image-to-image translation with conditional adversarial networks. IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Honolulu, 2017), pp. 5967–5976

42. A.B.L. Larsen, S.K. Sønderby, H. Larochelle, O. Winther, in *Autoencoding beyond pixels using a learned similarity metric. International Conference on Machine Learning* (ICML, New York, 2016) pp. 1558-1566
43. TS Yoo, *Insight Into Images: Principles and Practice for Segmentation, Registration and Image Analysis*, 1st edn. (A.K. Peters Ltd., Natick, 2004).
44. J.A. Jensen, FIELD: a program for simulating ultrasound systems. *Med Biol Eng Comput* **34**, 351–353 (1996)
45. J.A. Jensen, N.B. Svendsen, Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE Trans Ultrason Ferroelectr Freq Control* **39**(2), 262–267 (1992)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.