Democratizing LED-Based Photoacoustic Imaging with Adaptive Beamforming and Deep Convolutional Neural Network

Haichong K. Zhang

Abstract The current standard photoacoustic (PA) imaging technology includes two hardware requirements: high power pulsed laser for light illumination and multi-channel data acquisition device for PA signal recording. These requirements have been limiting factors to democratize PA imaging because a laser is heavy, expensive and includes hazardous risk, and most parallel data acquisition technology is available only in specialized research systems. The goal of this chapter is to provide an overview of technologies that will enable safer and more accessible PA imaging, as well as introduce the use of safe and fast light emitting diode (LED) light sources in combination with clinical ultrasound machines. There are two limiting factors that prevent achieving this. First, clinical ultrasound machines typically only provide post-beamformed data based on an ultrasound delay function, which is not suitable for PA reconstruction. Second, a PA image based on the LED light source suffers from low signal-to-noise-ratio due to limited LED-power and requires a large number of averaging. To resolve these challenges, an adaptive synthetic aperture beamforming algorithm is applied to treat defocused data as a set of pre-beamformed data for PA reconstruction. An approach based on deep convolutional neural network trains and optimizes the network to enhance the SNR of low SNR images by guiding its feature extraction at different layers of the architecture. We will review and discuss these technologies that could be the key to advancing LED based PA imaging to become more accessible and easier to translate into clinical applications.

H. K. Zhang (✉)
Department of Biomedical Engineering,
Worcester Polytechnic Institute, Worcester, MA 01609, USA
e-mail: hzhang10@wpi.edu

Robotics Engineering Program,
Worcester Polytechnic Institute, Worcester, MA 01609, USA

© Springer Nature Singapore Pte Ltd. 2020

M. Kuniyil Ajith Singh (ed.), *LED-Based Photoacoustic Imaging*,
Progress in Optical Science and Photonics 7,
https://doi.org/10.1007/978-981-15-3984-8_8

1 Introduction

Photoacoustic imaging is an emerging modality offering unique contrast of optical absorption and imaging depth of ultrasound for a wide range of biomedical applications [1]. The clinical accessibility of photoacoustic (PA) imaging is limited because of specific hardware requirements including high energy pulsed laser and channel data acquisition system [2–5]. Most laser systems used for PA imaging provide high pulse energy in the mJ scale with a low pulse repetition frequency (PRF) of 10–20 Hz. These laser systems are bulky, expensive, and unsafe, requiring eye protection, such as laser glasses. Installation at a hospital would require a special room that meets the laser safety requirements.

To democratize PA imaging toward broader clinical applications and its usage in research, a light source that is compact, low-cost, and safe to use is desired. Light emitting diode (LED) light sources have been considered as a viable alternative [6, 7]. Compared to high power laser systems, the LED-based light source has the advantage in terms of size, cost, and safety. Most importantly, the LED light source is not classified as a laser, so laser safety regulations such as the light shield and laser safety glasses are not required. The limitation of an LED light source is its low output power. Series of LEDs can generate energy only in the range of μJ, while common high-power pulsed laser used for PA imaging produce energy in the mJ range. Due to the low power output, the received PA signal of an LED-based system suffers from low signal-to-noise-ratio (SNR). Current technology aiming to improve the SNR is based on acquiring multiple frames of PA signals, and subsequently perform an averaging over them to minimize the noise. Though the pulse repetition frequency of a LED-based system is much higher (in range of kHz) than the high-power laser, an averaging over many frames, typically thousands, reduces the effective frame rate of PA images. Furthermore, a large number of averaging frames require longer scanning times, leading to potential motion artifacts in reconstructed PA images.

The other hardware challenge is the accessibility of data acquisition (DAQ) devices used for PA imaging [8, 9]. Pre-beamformed channel data from acquisition devices are required to collect the raw PA signals because PA reconstruction requires a delay function calculated based on the time-of-flight (TOF) from the light source to the receiving probe element, while US beamforming considers the round trip initiated from the transmitting and receiving probe element. Thus, the reconstructed PA image with an ultrasound beamformer would be defocused due to the incorrect delay function. Real-time channel data acquisition systems are only accessible from limited research platforms. Most of them are not FDA approved, which hinders the development of PA imaging in the clinical setting. Therefore, there is a demand to implement PA imaging on more widely used clinical machines.

To broaden the impact of clinical PA imaging, this paper presents a vendor-independent PA imaging system utilizing ultrasound post-beamformed data, which is readily accessible in some clinical scanners. While a LED light source with low energy output and high PRF is used to replace a conventional high energy laser, a deep neural networks-based approach is presented to improve the quality of PA

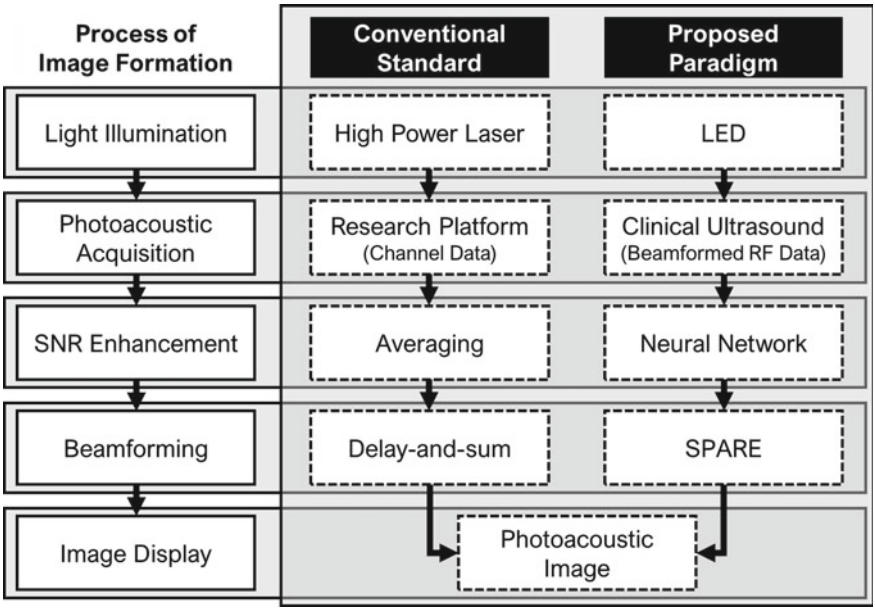

Fig. 1 The conventional photoacoustic imaging architecture and the new paradigm introduced in this chapter using LED light source and clinical ultrasound machine

images as well as reduce the number of averaging frames in image reconstruction. Figure 1 summarizes the process of PA image formation based on conventional architecture compared with the proposed paradigm incorporating a LED light source and a clinical ultrasound machine.

In this chapter, we review two enabling technologies for a LED-based and PA imaging system integrated with clinical ultrasound scanners; the image reconstruction approach using a post-beamformed RF data and the deep neural network-based SNR enhancer.

2 Image Reconstruction from Post-beamformed RF Data

2.1 Problem Statement

The acquisition of channel information is crucial to form a PA image, since typical clinical ultrasonic machines only provide access to beamformed data with delay-and-sum [2, 8]. Accessing pre-beamformed channel data needs customized hardware and parallel beamforming software and is available for dedicated research ultrasound platforms, such as the Ultrasonix DAQ system [9]. In general, these systems are costly with fixed data transfer rates that prohibit high frame rate, real-time imaging [10].

More importantly, PA beamforming is not supported by most clinical ultrasound systems. Harrison et al. has suggested changing the speed of sound parameter of clinical ultrasound systems [11]. Software access to alter the sound speed is not prevalent, however, and the range for this change is restricted when available, making this choice inadequate for reconstruction of PA images. In addition, the applicability of this technique is restricted to linear arrays, because angled beams (e.g. as in curvilinear arrays) change beamformer geometry and the speed of sound. Thus, compensation cannot be made by merely altering the sound velocity. In contrast, several clinical and research ultrasound systems have post-beamformed radio frequency (RF) data readily available. The objective in this section is to devise a PA image reconstruction approach based on ultrasound RF data that the system has already beamformed. A synthetic aperture-based beamforming algorithm, named Synthetic-aperture based PhotoAcoustic RE-beamforming (SPARE), utilizes ultrasound post-beamformed RF data as the pre-beamformed data for PA beamforming [12, 13]. When receive focusing is applied in ultrasound beamforming, the focal point can be regarded as a virtual element [14–16] to form a set of pre-beamformed data for PA beamforming. The SPARE beamformer takes the ultrasound data as input and outputs a PA image with the correct focal delay applied.

2.2 Technical Approach

2.2.1 Ultrasound Beamforming

The difference between ultrasound and PA beamforming is the acoustic time-of-flight (ToF) and related delay function. The delay function in delay-and-sum beamforming is calculated from the distance between the receivers and the target in ultrasound image reconstruction [17]. The acoustic wave is first transmitted from the ultrasound transducer via a medium with a specific velocity, reflected at boundaries, and the backscattered sound is received by the ultrasound transducer. The acoustic ToF during this process can be formulated as,

$$t_{US}(r_F) = \frac{1}{c}(|r_T| + |r_R|), \tag{1}$$

where r_F is the focus point originating from the ultrasound image coordinates, r_T is the vector from the transmit element to the focal point, r_R is the vector from the focal point to the receive element, and c is the speed of sound. Sequential beamforming with dynamic focus or fixed focus is applied as a delay-and-sum algorithm in clinical ultrasound systems. In dynamic focusing, the axial component, z_F, of the focusing point differs with depth, while a single fixed depth focus is used for the fixed focusing.

The acoustic TOF of PA signals is half of that of ultrasound, because the acoustic wave is produced at the target by absorbing light energy, and the time to travel from the

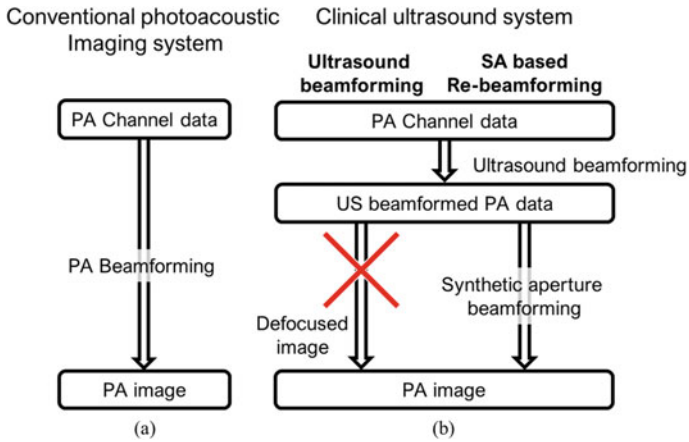

Fig. 2 Conventional PA imaging system (**a**) and proposed PA imaging system using clinical ultrasound scanners (**b**). Channel data is required for PA beamforming because ultrasound beamformed data is defocused with the incorrect delay function, where the introduced approach treats this information as pre-beamformed data for additional beamforming

optical transmission side negligible. Therefore, the acoustic TOF for photoacoustic imaging is

$$t_{PA}(r_F) = \frac{|r_R|}{c}. \tag{2}$$

Considering the differences between Eqs. (1) and (2), when beamforming is applied to the received PA signals based on Eq. (2), the beamformed RF signals are defocused (Fig. 2).

2.2.2 Synthetic Aperture-Based Re-beamforming

In the SPARE beamforming, the beamformed RF data from the ultrasound scanner is not considered as defocused useless data, but as pre-beamformed RF data for PA beamforming. The additional delay-and-sum step is applied on the beamformed RF data, and it is possible to reconstruct the new photoacoustically beamformed RF data. The focus point in the axial direction is constant with depth when fixed focusing is applied in the ultrasound beamforming process, suggesting that optimal focusing has been implemented at the particular focal depth with defocused signals appearing elsewhere. Initiating from the single focal depth, the defocused signals appear as if they were transmitted from the focal point (i.e. a virtual element as illustrated in Fig. 3b). In this sense, the ultrasound post-beamformed RF data is considered as PA pre-beamformed RF data. The TOF from the virtual element, when a fixed focus at z_F is applied, becomes

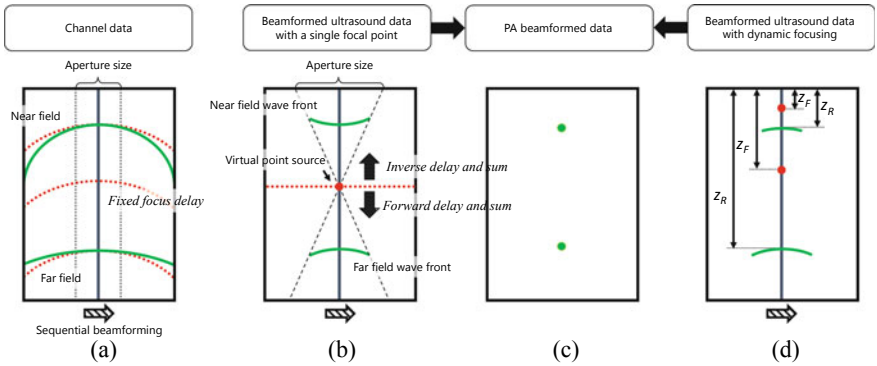

Fig. 3 Illustration of channel data and the SPARE-beamforming process [55]. **a** In channel data, the wave front of received RF signals expand corresponding to the depth (green line). The red lines indicate fixed focus delay function. **b** When fixed receive focusing is applied, the delay function is only optimized to the focus depth (red line). **c** As a result of fixed receive focusing, the focal point can be regarded as a virtual point source, so that inverse and forward delay and sum can be applied. **d** Similarly, dynamic focusing could be regarded as a specific case of that in which the virtual element depth z_F is the half distance of re-beamforming focal depth z_R

$$t\left(r'_F\right) = \frac{\left|r'_R\right|}{c}, \tag{3}$$

where

$$\left|r'_R\right| = \sqrt{(x_R)^2 + (z_R - z_F)^2}, \tag{4}$$

and $r'_F = r_F - z_F$. x_R and z_R is the lateral and axial components of r_R, respectively. The dynamic receive delay function is applied in the positive axial direction when $z_R \geq z_F$, and negative dynamic focusing delay is applied when $z_R < z_F$. The diagrams in Fig. 3b, c show the re-beamforming process of the SPARE-beamformer. Post-beamforming processes such as envelope detection and scan conversion are applied on the reconstructed data for the PA image display.

This theory applies to both fixed and dynamic focused beamformed ultrasound RF data with difference being that in dynamic focusing, the round-trip between the transmitter and the reflecting point in conventional ultrasound imaging must be considered along with the location of the virtual point source. Thus, in SPARE beamforming of dynamically focused data, the virtual point source depth, z_F, is considered to be dynamically varied by half of the photoacoustic beamforming focal point depth, z_R, as illustrated in Fig. 3d. Note that $z_R = 2z_F$ is always true in this special case.

2.3 Simulation Evaluation

The concept validation was performed through the ultrasound simulation tool, Field II [18]. A 128-element, 0.3 mm pitch, linear array transducer was assumed to be a receiver, which matches the setup of the experiment presented in Sect. 2.4. The standard delay-and-sum PA beamforming algorithm was applied to the simulated channel data in order to provide a ground-truth resolution value for this setup. Five-point targets were placed at depths of 10 mm to 50 mm with 10 mm intervals. To simulate defocused data, delay-and-sum with dynamic receive focusing and an aperture size of 4.8 mm was used to beamform the simulated channel data assuming ultrasound delays. The simulation results are shown in Fig. 4. The ultrasound beamformed RF data was defocused due to an incorrect delay function (Fig. 4b). The reconstructed PA images are shown in Figs. 4c–d. The measured full width at half maximum (FWHM) is shown in Table 1. The reconstructed point size was comparable to the point reconstructed using a 9.6 mm aperture on the conventional PA beamforming.

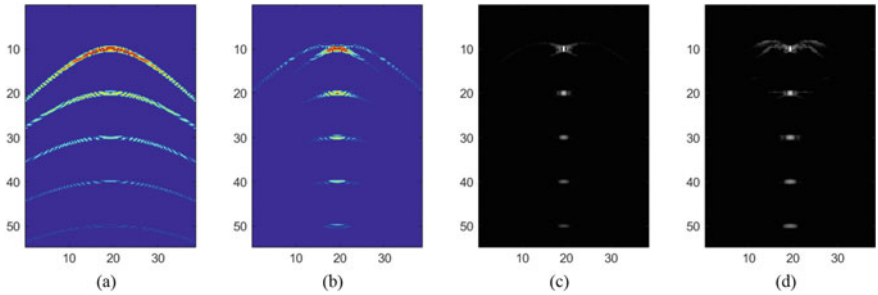

Fig. 4 Simulation results. **a** Channel data. **b** Ultrasound post-beamformed RF data. **c** Reconstructed PA image from channel data with an aperture size of 9.6 mm. **d** Reconstructed PA image through SPARE beamforming

Table 1 FWHM of the simulated point targets for corresponding beamforming methods

FWHM (mm)	Control using channel data	SPARE-beamforming
10 mm depth	0.60	0.63
10 mm depth	1.02	0.99
10 mm depth	1.53	1.43
10 mm depth	1.94	1.91
10 mm depth	2.45	2.42

2.4 Experimental Demonstration

The PA sensing system was employed for evaluating the LED-based PA imaging performance; a near-infrared pulsed LED illumination system (CYBERDYNE INC, Tsukuba, Japan) was used for PA signal generation. To collect the generated PA signals, a clinical ultrasound machine (Sonix Touch, Ultrasonix) with a 10 MHz linear ultrasound probe (L14-5/38, Ultrasonix) was used to display and save the received data. A line phantom made with fishing wire was imaged to evaluate the SNR and resolution performance. The ultrasound post-beamformed RF data with dynamic receive focusing was then saved. To validate the channel data recovery through inverse beamforming, the raw channel data was collected using a data acquisition device (DAQ). Figure 5 shows the experimental results imaging the cross section of a line phantom [19, 20]. The control data was reconstructed from channel data collected from the DAQ. The SPARE result used the ultrasound post-beamformed data collected from the ultrasound scanner as the input. The SPARE algorithm produced better imaging contrast and SNR when comparing the inherent resolution of the two methods. By quantifying the SNR change over the number of averaging, these two were correlated in a log-linear model for both with and without the use of channel data, depicted in Fig. 5b. In result, the gradient of the SPARE method was larger than conventional PA reconstruction from channel data, because the ultrasound beamformed data was summed already across the aperture once even with incorrect focus, and the random

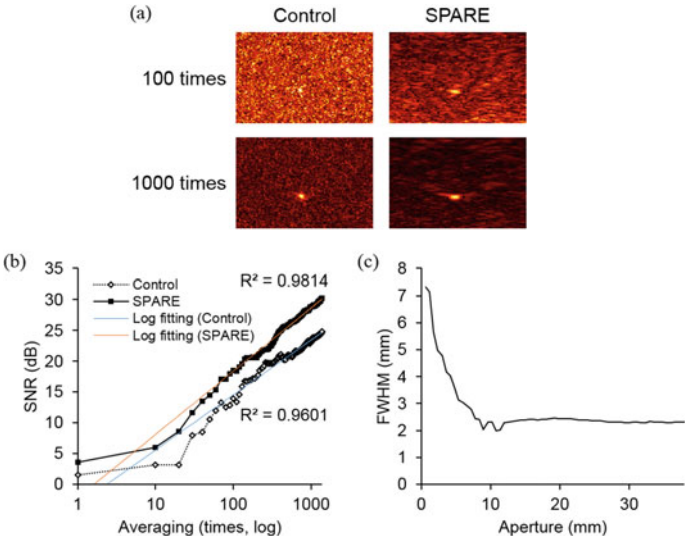

Fig. 5 Experiment results with LED light source imaging line phantom. **a** Comparison of control using channel data from DAQ and SPARE results using ultrasound post-beamformed data. **b** SNR analysis of both control and SPARE results. **c** Resolution analysis of SPARE results. Resolution improvement was hindered at FWHM of 2 mm due to the aperture size [19, 20]

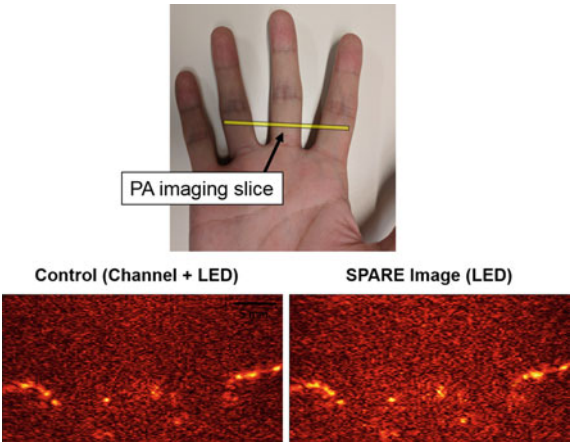

Fig. 6 In vivo PA imaging of human fingers using LED light source. Experimental configuration of ultrasound and PA images of human fingers are shown. PA images were reconstructed using channel data from a DAQ device and beamformed RF data with the SPARE algorithm [19, 20]

noise can be suppressed in this process. The control result showed better spatial resolution compared to the SPARE result because the ultrasound beamformed data was formed from a restricted aperture size (maximum 32 elements) due to restriction of the ultrasound scanner, while the channel data could utilize the complete aperture for reconstruction (Fig. 5c).

Human fingers were imaged using 850 nm LED bars for an in vivo experiment (Fig. 6). The channel data was collected first, then the ultrasound beamformed data was produced to compare standard and suggested solutions to beamforming. The raw channel data was averaged 3000 times to maximize the imaging contrast. It was verified that the SPARE approach could achieve comparable image quality to the channel data.

2.5 Discussion

The introduced SPARE method would work for any structures that have high optical absorption such as blood vessels that show strong contrast for near-infrared wavelength light excitation. Reconstruction artifacts such as side lobe and grating lobe could appear and influence non-point targets making the image quality of SPARE image was worse than standard PA image using channel data. The algorithm could also be incorporated with clinical ultrasound machines in real-time imaging schemes. Theoretically, the SNR of two beamformers should be similar, and this discrepancy could be attributed to the summation of axially distributed coherent information twice, once for each beamforming step. When the SNR of the channel signals is considerably low, the reconstructed image may contain a noise-related gradation artifact as the number of summations differs for each focal point. Hence, beamforming with the full aperture is more suitable in this high-noise situation. The image quality improvement strategies (apodization, transmit gain compensation, etc.) are

expected to have a comparable impact on the SPARE image enhancement. Apodization improves the appearance of the reconstructed image, because it reduces the sidelobes in the ultrasound beam.

The suggested technique is superior than the speed of sound adjustment approach [11] and is applicable to steered beams (e.g. phased arrays) and to beam geometries that vary from linear arrays (e.g. curvilinear arrays). As formulated in Eqs. (3) and (4), the proposed beamformer applies a delay-and-sum assuming the PA signals are received at the virtual element. Therefore, even if the ultrasound beam is angled, the delay-and-sum algorithm is still applicable with the virtual element created by the angled beam.

Suppression of ultrasound transmission may be regarded as another system requirement. The ideal solution is to turn off the transmit events. However, if this function is not available, an option is to lower the transmission energy voltage. The use of an electrical circuit to regulate the timing of the laser transmission is another strategy. Subtracting the images with and without laser excitation would highlight the PA signals.

One system requirement for the SPARE beamformer is a high pulse repetition frequency (PRF) laser system. In order to maintain the frame rate, so that it is comparable to that of ultrasound B-mode imaging, the PRF of the laser transmission should be the same as the ultrasound transmission rate, in the range of at least several kHz. In fact, a high PRF laser system, such as a LED, is idealistic. Based on the assumption that the LED frame rate is 16,000 and the reception ultrasound has 128 lines of acquisition, Fig. 7 summarizes the estimated frame rate and laser energy by varying the number of averaging. Since SNR improvement under averaging is the square root of the number of averaging, outputting 1 mJ and 5 mJ light source energy requires 25 and 625 times averaging, respectively. The highest frame rate available when the DAQ unit is accessible is 625 and 25.6 frames per second, respectively. When a clinical ultrasound scanner was used for data acquisition, the frame rate becomes 5 and 0.2 frames per second, respectively. Using clinical ultrasound machine, the highest frame rate available is 125 without averaging.

The novelty of the SPARE algorithm suggested its potential for integration with clinical ultrasound scanners to become real-time imaging systems [21]. Most real-time photoacoustic imaging systems are currently based on open platform research systems [9]. However, the option of using a clinical ultrasound system already with FDA approval eases the transition of photoacoustic technology into the clinic. Potential applications include in vivo real-time photoacoustic visualization for brachytherapy monitoring [22–24], brain imaging [25–28], image-guided surgery [29, 30], interventional photoacoustic tracking [31], multispectral interventional imaging [32, 33], and cardiac radiofrequency ablation monitoring [34].

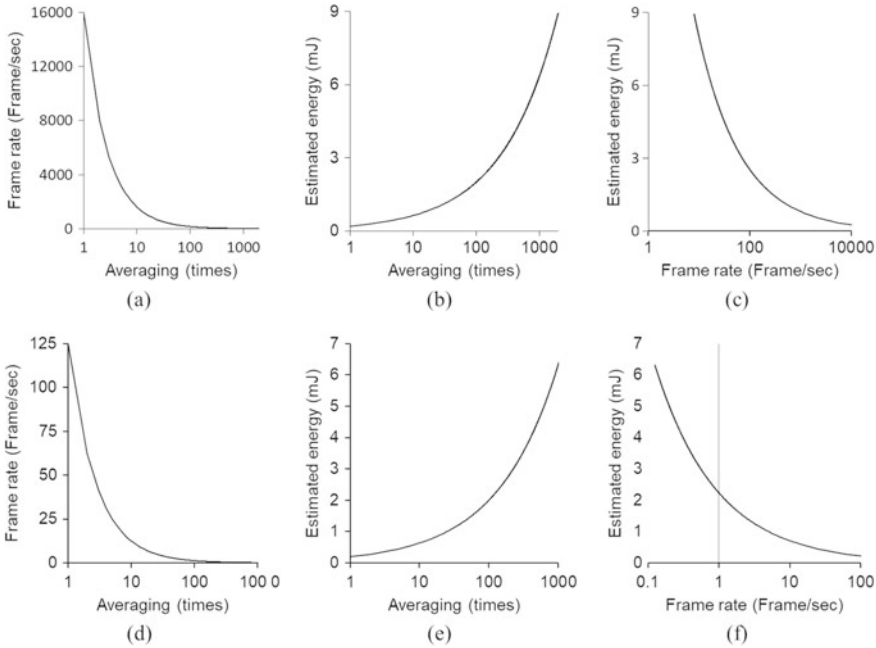

Fig. 7 Numerical estimation of frame rate using a LED system. Frame rate (**a**, **d**) and estimated energy (**b**, **e**) by varying the number of averaging, and the relationship between frame rate and estimated energy (**c**, **f**) are shown using a DAQ device (**a**–**c**), and using a clinical ultrasound system (**d**–**f**) for data collection

3 SNR Enhancement with Convolutional Neural Network

3.1 Problem Statement

The most classic and conventional strategy to improve the SNR with a low-power light source such as the LED-based scheme is averaging, obtaining multiple frames (ten, hundreds, or a few thousand) of the same sample, then averaging them over. When the noise has its distribution of ɔ, the noise distortion after the averaging of N times is expressed as

$$\sigma_{avg-N} = \frac{\sqrt{N}}{N}\sigma, \tag{5}$$

and the SNR improvement is proportional to the number of frames used for averaging. While using more frames to average earns an enhanced SNR, it decreases PA imaging's effective frame rate. Reduced frame rate makes it difficult to adapt this technology to moving objects, like the heart, and prone to motion artifacts. The signal processing approaches, such as adaptive denoising, empirical mode decomposition,

wavelet transform or Wiener deconvolution could be used to tackle the limitation of averaging [7, 35]. Coded excitation is a strategy that increases the SNR without compromising the measurement time. In temporal encoding, the laser pulses are sent with a special encoded pattern without the need for waiting the acoustic TOF. The PA signals with an improved SNR are decoded from the received encoded RF signals. Golay codes [36] and m-sequence family (such as preferred pairs of m-sequences and Gold codes) [37, 38] have been proposed for temporal encoding. The limitation of coded excitation is that it presents its benefit only if the pulse interval is shorter than that of the acoustic TOF, thus ultra-high PRF lasers with hundreds kHz or several MHz pulsing capabilities are required. Therefore, a more generalized approach is needed to improve the SNR for the usage of LED light source.

A recently emerging approach based on deep convolutional neural networks is a powerful alternative. Deep neural networks have been introduced to image classification [39, 40], image segmentation [41], image denoising [42] and image super-resolution [43–46] and outperforms state-of-the-art signal processing approaches. The published image enhancement techniques are based on stacked denoising auto-encoder [42], densely connected convolutional net [46] or including perceptual loss to enhance the spatial structure of images [44]. Neural networks have been applied on PA imaging for image reconstruction [47–49] and removal of reflection artifacts [50]. This section focuses on the usage of deep convolutional neural network to differentiate the main signal from the background noise and to denoise a PA image with a reduced number of averaging.

The introduced architecture consists of two key components; one is convolutional neural networks (CNN) that extracts the spatial features, and the other one is recurrent neural networks (RNN) that leverages the temporal information in PA images. The CNN is built upon a state-of-the-art dense net-based architecture [46] that uses series of skip-connections to enhance the image content. Convolutional variant of short-long-term-memory [51, 52] is used for the RNN to exploit the temporal dependencies in a given PA image sequence. Skip-connections are integrated throughout the networks, including both CNN and RNN components, to effectively propagate features and eliminate vanishing gradients. While the full description of approaches can be found in Refs. [53, 54], this section provides digest of them.

3.2 Deep Convolutional Neural Network

A dense net-based CNN architecture to denoise PA images is introduced by Anas et al. [46, 53, 54]. The PA image with a limited number of averaging is used as the input, and the objective is to produce a high-quality PA image that provides an equivalent SNR compared to a PA image with a considerably high number of averaging. Figure 8 shows the deep neural network architecture [46]. The network focusing on improving the image quality of a single PA image is illustrated in Fig. 8a. The number of feature maps in each convolutional layer is defined as ‘xx’ in ‘Conv xx’. The architecture consists of three dense blocks, and each dense block is composed of two densely

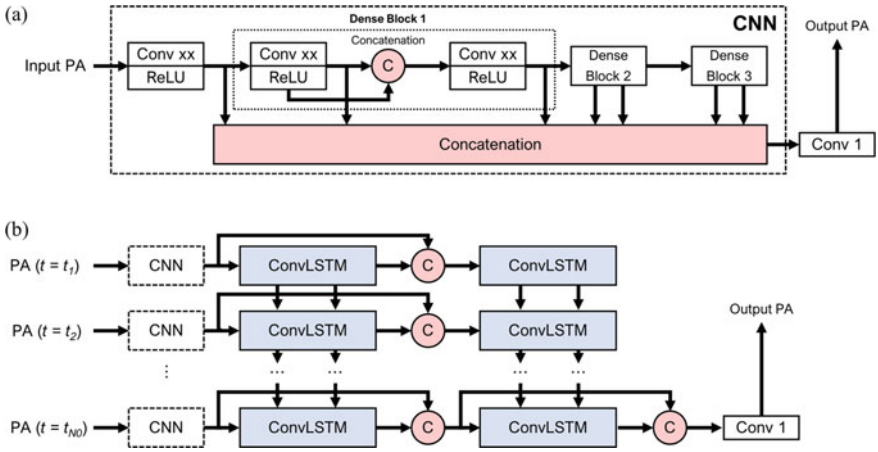

Fig. 8 A schematic of the introduced deep neural network-based approach (Reproduced from [53]). **a** The dense net-based CNN architecture to improve the quality of PA image. The architecture consists of three dense blocks, each dense block includes two 3 by 3 dense convolutional layers followed by rectified linear units. **b** The architecture that integrates CNN and ConvLSTM together to extract the spatial features and the temporal dependencies, respectively

connected convolutional layers and rectified linear units (ReLU). The benefit of using the dense convolutional layer is elimination of the vanishing gradient problem of deep networks [55] because all the features initially produced are inherited and succeeded in the following layers. The output image is produced by convoluting the feature map with all features from the concatenated dense blocks.

In addition to CNN, a recurrent neural network (RNN) [56, 57] is implemented to mitigate the temporal dependencies in a specific sequence. While several variants of RNN have been reported, and long-short-term-memory (LSTM) [51] showed the most successful performance in different applications. ConvLSTM [52] is an extension of LSTM that uses the convolution operation to extract temporal features from a series of 2D maps. The introduced architecture combining CNN and ConvLSTM to improve the denoising performance is shown in Fig. 8b. The architecture takes as inputs a series of PA images in different time points. It initially uses CNN to obtain the spatial features and then subsequently utilizes ConvLSTM to exploit the temporal dependencies. Two layers of ConvLSTM including skip connections are used for the recurrent connection. At the end, all the features generated in the previous layers are concatenated to compute the SNR improved PA image as the final output.

3.3 Experimental Demonstration

The concept was validated by training the network and assessing the SNR enhancement with a point target and proved further with human fingers in vivo. Two sets of

LED bar-type illuminators were placed on both sides of a linear ultrasound transducer array for the image setup. The LED's pulse repetition frequency was set at 1 kHz and PA data acquisition was synchronized with the LED excitation. PA images of the point target from the wire phantom were used to train the neural networks, assuming that those PA images with multiple point targets at different depths enable our network to learn how to improve the quality of the point spread function. The trained network with a point spread function can be applied to any arbitrary function of PA target.

The number of averaging was used to control the reconstructed image quality to produce input data consisting of low and high SNR target PA images for the training. For low SNR inputs, lower values of N in the range of 200–11,000 was chosen, with a step of 200. The averaging frame numbers in the sequence can be represented as $\{N_s; 2N_s; 3N_s; \ldots; N_0\}$ corresponding to time index $\{t_1; t_2; t_3; \ldots; t_{N0}\}$, where N_s was set to 200, and N_0 was 11,000. For each chosen value of N, the large set of 11,000 frames was split into several subsets, where each subset consists of N frames of PA signals. For each subset of N frames data, the PA signals are averaged first, followed by reconstruction to obtain one post-processed PA image. With the collection of 11,000 frames for one phantom sample, the greatest possible quality PA image can be achieved by reconstructing it from the averaged signal over all frames, which is regarded as the ground truth target image. Note that for each experiment, there is only one gold-standard target image that corresponds to more than one input sequences. Mean square losses are used as a loss function between the predicted and gold-standard target PA images. To minimize the loss function, TensorFlow library (Google, Mountain View, CA) with Adam [58] optimization technique is used. The quantitative assessment was performed with the independent test dataset. The peak-signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used as evaluation indices that compare the output of our networks with the highest quality target image [59].

The comparison of PSNR and SSIM of two techniques using deep neural networks (CNN-only and RNN + CNN) for different averaging frame numbers is shown in Fig. 9a, b. The solid line in the figure shows the mean value for each computing method calculated from 30 test samples. The shaded region reflects the corresponding standard deviation of each evaluation index. While both deep neural network approaches outperform the SNR enhancement over averaging, the approach of RNN + CNN presented the highest performance among them. The improvements of RNN + CNN in PSNRs of 5.9 dB and 2.9 dB was accomplished on average with respect to averaging and CNN-only techniques, respectively. Figure 9c presents the amount of frame rate enhancement two deep neural network approaches relative to averaging to attain certain PSNA. The gain is calculated with respect to the frame number of the averaging approach. For example, at a mean PSNR of 35.4 dB, the RNN + CNN, CNN-only and averaging techniques need 1360, 3680 and 11,000 averaging frames, respectively. When the averaging approach was treated as reference, the RNN + CNN and CNN-only achieved gains in the frame rate of 8.1 and 3.0 times, respectively. With the deep neural network approaches, the improved frame rate can be achieved without compromising the SNR.

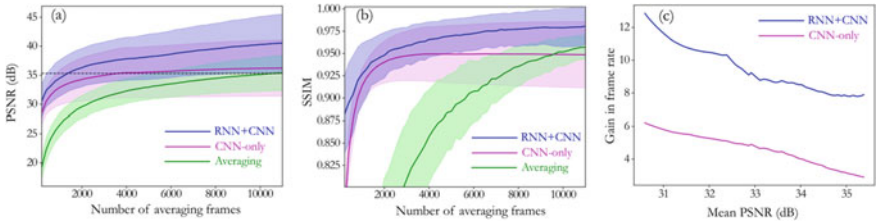

Fig. 9 A comparison of PSNR and SSIM of our (RNN + CNN) method with those from the simple averaging and CNN-only methods [53]. **a** PSNR versus averaging frame numbers. An improvement at all the averaging frame numbers is seen for our method compared to the two other methods. A higher improvement rate of the method is observed compared to the CNN-only method. **b** SSIM versus averaging frame numbers. Unlike CNN-only method, the trend of improvement is observed with the averaging frame numbers for our method. **c** Gain in frame rate versus mean PSNR

Figure 10 shows a qualitative comparison among all three comparative methods for proper digital arteries of three fingers of a volunteer (anatomy is shown at bottom in the figure). Three blood veins were noticeable for each finger, where enhanced blood vessel detections were observed for the RNN + CNN approach (highlighted by arrows). Note that the PA image averaged from the 5000 frames (high quality in the figure) includes some remaining noises and artifacts due to the movement during the scanning period.

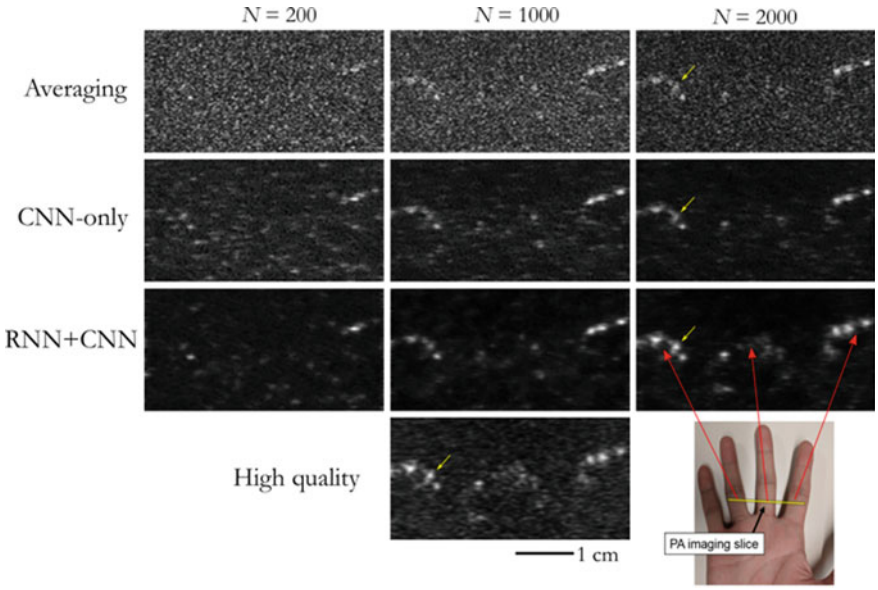

Fig. 10 A comparison of our method with the averaging and CNN-only techniques for an in vivo example [53]. Improvements are noticeable compared to those of other two methods in recovering the blood vessels (marked by arrows)

The GPU computation times are 15 and 4 ms for RNN + CNN and CNN-only methods, respectively. The corresponding run-times in the CPU are 190 and 105 ms, respectively.

3.4 Discussion

This section presented a deep neural networks approach to improve the quality of PA images in real-time while simultaneously reducing the scanning period. Besides using CNN to obtain the spatial features, RNN is used in the architecture to exploit the temporal information in PA images. The network was trained using a sequence of PA images from 32 phantom experiments. On the test from 30 samples, a gain in the frame rate of 8.1 times is achieved with a mean PSNR of 35.4 dB compared to the conventional averaging approach. A temporal PA sequence allows the neural networks to learn the image and noise contents more effectively than a single image-based CNN-only network does. In addition, for the CNN-only method, saturation in both image quality indices is observed for higher averaging frame numbers (Fig. 9a, b) indicating a decrease in the rate of improvement with a rise in the averaging frame number, as opposed to the higher improvement rate for the CNN + RNN method. Furthermore, the improved performance of the deep neural network approach was demonstrated through an in vivo example (Fig. 10). The key benefit of the technique is that it could improve the image quality from a reconstructed image with low averaging frame number, thus eliminating the potential effect of the artifacts.

4 Conclusions and Future Directions

In this chapter, we reviewed a paradigm on PA imaging using LED light source and image reconstruction with ultrasound post-beamformed RF data from a clinical ultrasound system. SPARE-beamforming takes the post-beamformed data and compensate the delay error by producing a PA image. Simulation and experimental studies presented that this approach can achieve an equivalent resolution compared to PA image generated from channel data. In addition, it was demonstrated that deep neural networks have a potential to exploit the temporal information in PA images for an improvement in image quality as well as a gain in the imaging frame rate.

Future directions along the line of this research include the exploration of developing beamforming algorithms utilizing more accessible data from clinical ultrasound machine such as post-beamformed B-mode images. It was reported that if the PA target is a point-like target, the post-beamformed B-mode image can be used as the source for PA image recovery [60]. For the image quality enhancement based on deep neural networks, more extensive in vivo evaluations are required to validate the clinical translatability. Other training architectures that do not take the final B-mode

image but pre-beamformed or post-beamformed RF data may enhance both SNR and resolution of a PA image [61].

Acknowledgments The author would like to acknowledge the Medical UltraSound Imaging and Intervention Collaboration (MUSiiC) laboratory led by Dr. Emad Boctor for providing the opportunity to develop the technologies introduced in this book chapter. The author also thanks Mr. Doua Vang for proof-reading the manuscript and providing helpful comments. Financial support was provided by the Worcester Polytechnic Institute internal fund and the National Institute of Health grant DP5 OD0Z8162 for preparing the manuscript.

Compliance with Ethical Standards

Conflict of Interest All authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed Consent Additional informed consent to publish was obtained from all individual participants included in the study.

References

1. M. Xu, L.V. Wang, Photoacoustic imaging in biomedicine. Rev. Sci. Instrum. **77**, 041101 (2006)
2. S. Park, S.R. Aglyamov, S. Emelianov, Beamforming for photoacoustic imaging using linear array transducer, in *Proceedings of IEEE International Ultrasonics Symposium* (2007), pp. 856–859
3. B. Yin, D. Xing, Y. Wang, Y. Zeng, Y. Tan, Q. Chen, Fast photoacoustic imaging system based on 320-element linear transducer array. Phys. Med. Biol. **49**(7), 1339–1346 (2004)
4. C.K. Liao, M.L. Li, P.C. Li, Optoacoustic imaging with synthetic aperture focusing and cohehrence weighting. Opt. Lett. **29**, 2506–2508 (2004)
5. R.G.M. Kolkman, P.J. Brands, W. Steenbergen, T.G.V. Leeuwen, Real-time in vivo photoacoustic and ultrasound imaging. J. Biomed. Opt. **13**(5), 050510 (2008)
6. A. Hariri, J. Lemaster, J. Wang, A.S. Jeevarathinam, D.L. Chao, J.V. Jokerst, The characterization of an economic and portable LED-based photoacoustic imaging system to facilitate molecular imaging. Photoacoustics **9**, 10–20 (2018)
7. X. Dai, H. Yang, H. Jiang, In vivo photoacoustic imaging of vasculature with a low-cost miniature light emitting diode excitation. Opt. Lett. **42**(7), 1456–1459 (2017)
8. J.J. Niederhauser, M. Jaeger, M. Frenz, Comparision of laser-induced and classical ultrasound. Proc. SPIE **4960**, 118–123 (2003)
9. N. Kuo, H.J. Kang, D.Y. Song, J.U. Kang, E.M. Boctor, Real-time photoacoustic imaging of prostate brachytherapy seeds using a clinical ultrasound system. J. Biomed. Opt. **17**(6), 066005 (2012)
10. H.J. Kang, N. Kuo, X. Guo, D. Song, J.U. Kang, E.M. Boctor, Software framework of a real-time pre-beamformed RF data acquisition of an ultrasound research scanner. Proc. SPIE **8320**, 83201F (2012)
11. T. Harrison, R.J. Zemp, The applicability of ultrasound dynamic receive beamformers to photoacoustic imaging. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **58**(10), 2259–2263 (2011)

12. H.K. Zhang, M. Bell, X. Guo, H.J. Kang, E.M. Boctor, Synthetic-aperture based photoacoustic re-beamforming (SPARE) approach using beamformed ultrasound data. Biomed Opt Express **7**(8), 3056–3068 (2016)
13. H.K. Zhang, X. Guo, B. Tavakoli, E.M. Boctor, Photoacoustic imaging paradigm shift: towards using vendor-independent ultrasound scanners. Med. Image Comput. Comput.-Assist. Interv. (MICCAI) **9900**, 585–592 (2016)
14. C.H. Frazier, W. D. O'Brien, Synthetic aperture techniques with a virtual source element. IEEE Trans. Ultrason. Ferroelec. Freq. Contr. **45**, 196–207 (1998)
15. S.I. Nikolov, J.A. Jensen, Virtual ultrasound sources in high resolution ultrasound imaging. Proc. SPIE, Progr. Biomed. Opt. Imaging **3**, 395–405 (2002)
16. J. Kortbek, J.A. Jensen, and K. L. Gammelmark, Synthetic aperture sequential beamforming, in *Proceedings of IEEE International Ultrasonics Symposium* (2008), pp. 966–969
17. K.E. Thomenius, Evolution of ultrasound beamformers, in *Proceedings of IEEE Ultrasonics Symposium*, vol. 2 (1996), pp. 1615–1622
18. J. A. Jensen and N. B. Svendsen, Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers, in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39 (1992), pp. 262–267
19. H. Zhang (2017) Enabling technologies for co-robotic translational ultrasound and photoacoustic imaging. Doctoral dissertation, Johns Hopkins University
20. H.K. Zhang, J. Kang, E.M. Boctor, In vivo low-cost photoacoustic imaging using LED light source and clinical ultrasound scanner. Proc. SPIE (presentation), 10494-177 (2018)
21. H.K. Zhang, H. Huang, C. Lei, Y. Kim, E.M. Boctor, Software-based approach toward vendor independent real-time photoacoustic imaging using ultrasound beamformed data, in *Photons Plus Ultrasound: Imaging and Sensing 2017*, vol. 10064 (International Society for Optics and Photonics, 2017), p. 100643O
22. M.A.L. Bell, N.P. Kuo, D.Y. Song, J. Kang, E.M. Boctor, In vivo visualization of prostate brachytherapy seeds with photoacoustic imaging. J. Biomed. Opt. **19**(12), 126011 (2014)
23. M.A.L. Bell, X. Guo, D.Y. Song, E.M. Boctor, Transurethral light delivery for prostate photoacoustic imaging. J. Biomed. Opt. **20**(3), 036002 (2015)
24. H.K. Zhang, A. Cheng, E.M. Boctor, Three-dimensional photoacoustic imaging using robotically tracked transrectal ultrasound probe with elevational synthetic aperture focusing, in *Proceedings of CLEO* (2016)
25. X. Wang, Y. Pang, G. Ku, X. Xie, G. Stoica, L.V. Wang, Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain. Nat. Biotechnol. **21**, 803–806 (2003)
26. W. Li, R. Chen, J. Lv, H. Wang, Y. Liu, Y. Peng et al., In vivo photoacoustic imaging of brain injury and rehabilitation by high-efficient near-infrared dye labeled mesenchymal stem cells with enhanced brain barrier permeability. Adv. Sci. (Weinh) **5**, 1700277 (2018)
27. J. Kang, H.K. Zhang, S.D. Kadam, J. Fedorko, H. Valentine, A.P. Malla, P. Yan, M. Harraz, J.U. Kang, A. Rahmim, A. Gjedde, L.M. Loew, D.F. Wong, E.M. Boctor, Transcranial recording of electrophysiological neural activity in the rodent brain in vivo using functional photoacoustic imaging of near-infrared voltage-sensitive dye. Front. Neurosci. **13**, 579 (2019)
28. J. Kang, E.M. Boctor, S. Adams, E. Kulikowicz, H.K. Zhang, K. Raymond, E. Graham, Validation of noninvasive photoacoustic measurements of sagittal sinus oxyhemoglobin saturation in hypoxic neonatal piglets. J. Appl. Physiol. **125**(4), 983–989 (2018)
29. L. Xi, G. Zhou, N. Gao, L. Yang, D.A. Gonzalo, S.J. Hughes, H. Jiang, Photoacoustic and fluorescence image-guided surgery using a multifunctional targeted nanoprobe. Ann. Surg. Oncol. **21**(5), 1602–1609 (2014)
30. M.A.L. Bell, A.K. Ostrowski, K. Li, P. Kazanzides, E.M. Boctor, Localization of transcranial targets for photoacoustic-guided endonasal surgeries. Photoacoustics **3**(2), 78–87 (2015)
31. A. Cheng, H.J. Kang, H.K. Zhang, R.H. Taylor, E.M. Boctor, Ultrasound to video registration using a bi-plane transrectal probe with photoacoustic markers. Proc. SPIE **9786**, 97860J (2016)

32. W. Xia, D.I. Nikitichev, J. Mari, S.J. West, R. Pratt, A.L. David, S. Ourselin, P.C. Beard, A.E. Desjardins, Performance characteristics of an interventional multispectral photoacoustic imaging system for guiding minimally invasive procedures. J. Biomed. Opt. **20**(8), 086005 (2015)
33. H.K. Zhang, Y. Chen, J. Kang, A. Lisok, I. Minn, M.G. Pomper, E.M. Boctor, Prostate specific membrane antigen (PSMA)-targeted photoacoustic imaging of prostate cancer in vivo. J. Biophotonics, e201800021 (2018)
34. N. Dana, L.D. Biase, A. Natale, S. Emelianov, R. Bouchard, In vitro photoacoustic visualization of myocardial ablation lesions. Heart Rhythm **11**(1), 150–157 (2014)
35. S.A. Ermilov, T. Khamapirad, A. Conjusteau, M.H. Leonard, R. Lacewell, K. Mehta, T. Miller, A.A. Oraevsky, Laser optoacoustic imaging system for detection of breast cancer. J. Biomed. Opt. **14**(2), 024007 (2009)
36. M.P. Mienkina, C.-S. Friedrich, N.C. Gerhardt, M.F. Beckmann, M.F. Schiffner, M.R. Hofmann, G. Schmitz, Multispectral photoacoustic coded excitation imaging using unipolar orthogonal golay codes. Opt. Express **18**, 9076–9087 (2010)
37. H. Zhang, K. Kondo, M. Yamakawa, T. Shiina, Simultaneous multispectral coded excitation using gold codes for photoacoustic imaging. Jpn. J. Appl. Phys. **51**, 07GF03 (2012)
38. H.K. Zhang, K. Kondo, M. Yamakawa, T. Shiina, Coded excitation using periodic and unipolar M-sequences for photoacoustic imaging and flow measurement. Opt. Express **24**, 17–29 (2016)
39. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105
40. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *IEEE CVPR* (2016), pp. 770–778
41. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 3431–3440
42. J. Xie, L. Xu, E. Chen, Image denoising and inpainting with deep neural networks, in *Advances in Neural Information Processing Systems* (2012), pp. 341–349
43. C. Dong, C.C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. **38**, 295–307 (2016)
44. J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in *European Conference on Computer Vision* (Springer, 2016), pp. 694–711
45. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang et al., Photo-realistic single image super-resolution using a generative adversarial network, arXiv preprint (2016)
46. T. Tong, G. Li, X. Liu, Q. Gao, Image super-resolution using dense skip connections, in *2017 IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017), pp. 4809–4817
47. S. Antholzer, M. Haltmeier, J. Schwab, Deep learning for photoacoustic tomography from sparse data. arXiv preprint arXiv:1704.04587 (2017)
48. J. Schwab, S. Antholzer, R. Nuster, M. Haltmeier, DALnet: high-resolution photoacoustic projection imaging using deep learning. arXiv preprint arXiv:1801.06693 (2018)
49. S. Antholzer, M. Haltmeier, R. Nuster, J. Schwab, Photoacoustic image reconstruction via deep learning, in *Photons Plus Ultrasound: Imaging and Sensing 2018*, vol. 10494 (International Society for Optics and Photonics, 2018), p. 104944U
50. D. Allman, A. Reiter, M.A.L. Bell, Photoacoustic source detection and reflection artifact removal enabled by deep learning. IEEE Trans. Med. Imaging **37**(6), 1464–1477 (2018)
51. S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. **9**, 1735–1780 (1997)
52. S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, Convolutional LSTM network: a machine learning approach for precipitation nowcasting, in *Advances in Neural Information Processing Systems* (2015), pp. 802–810.
53. E.M.A. Anas, H.K. Zhang, J. Kang, E.M. Boctor, Enabling fast and high quality LED photoacoustic imaging: a recurrent neural networks based approach. Biomed. Opt. Express **9**(8), 3852 (2018)

54. E.M.A. Anas, H.K. Zhang, J. Kang, E.M. Boctor, Towards a fast and safe LED-based photoacoustic imaging using a deep convolutional neural networks. MICCAI (2018)
55. G. Huang, Z. Liu, K.Q. Weinberger, L. van der Maaten, Densely connected convolutional networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1 (2017), p. 3
56. J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, in *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* (World Scientific, 1987), pp. 411–415
57. D.E. Rumelhart, G.E. Hinton, R.J. Williams et al., Learning representations by back-propagating errors. Cogn. Model. **5**, 1 (1988)
58. D. Kingma, J. Ba, Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
59. Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**, 600–612 (2004)
60. H.K. Zhang, X. Guo, H.J. Kang, E.M. Boctor, Photoacoustic image reconstruction from ultrasound post-beamformed B-mode image. Proc. SPIE Photonics West BiOS 2016 **9708**(970837) (2016)
61. E.M.A. Anas, H.K. Zhang, C. Audigier, E.M. Boctor, Robust photoacoustic beamforming using dense convolutional neural networks. In *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation* (Springer, Cham 2018), pp. 3–11