Reconstruction of initial pressure from limited view photoacoustic images using deep learning

Dominik Waibel, Janek Gröhl, Fabian Isensee, Thomas Kirchner, Klaus Maier-Hein, et al.

Dominik Waibel, Janek Gröhl, Fabian Isensee, Thomas Kirchner, Klaus Maier-Hein, Lena Maier-Hein, "Reconstruction of initial pressure from limited view photoacoustic images using deep learning," Proc. SPIE 10494, Photons Plus Ultrasound: Imaging and Sensing 2018, 104942S (19 February 2018); doi: 10.1117/12.2288353

SPIE.

Event: SPIE BiOS, 2018, San Francisco, California, United States

Reconstruction of initial pressure from limited view photoacoustic images using deep learning

Dominik Waibel[a], Janek Gröhl[a,d], Fabian Isensee[b], Thomas Kirchner[a,c], Klaus Maier-Hein[b,d], and Lena Maier-Hein[a,d]

[a]Division of Computer Assisted Medical Interventions (CAMI), German Cancer Research Center (DKFZ), Heidelberg, Germany
[b]Division of Medical Image Computing (MIC), German Cancer Research Center (DKFZ), Heidelberg, Germany
[c]Department of Physics and Astronomy, Heidelberg University, Germany
[d]Medical Faculty, Heidelberg University, Germany

ABSTRACT

Quantification of tissue properties with photoacoustic (PA) imaging typically requires a highly accurate representation of the initial pressure distribution in tissue. Almost all PA scanners reconstruct the PA image only from a partial scan of the emitted sound waves. Especially handheld devices, which have become increasingly popular due to their versatility and ease of use, only provide limited view data because of their geometry. Owing to such limitations in hardware as well as to the acoustic attenuation in tissue, state-of-the-art reconstruction methods deliver only approximations of the initial pressure distribution. To overcome the limited view problem, we present a machine learning-based approach to the reconstruction of initial pressure from limited view PA data. Our method involves a fully convolutional deep neural network based on a U-Net-like architecture with pixel-wise regression loss on the acquired PA images. It is trained and validated on *in silico* data generated with Monte Carlo simulations. In an initial study we found an increase in accuracy over the state-of-the-art when reconstructing simulated linear-array scans of blood vessels.

Keywords: Photoacoustics, Machine learning, Deep learning, Photoacoustic tomography, Limited view problem

1. INTRODUCTION

Photoacoustic imaging (PAI) is a novel and fast emerging modality that combines the strengths of ultrasound (US) and optical imaging and offers high optical contrast as well as great imaging depth and spatial resolution.[1] In multispectral applications, PAI is especially suitable for real-time imaging of functional tissue parameters such as local blood oxygenation or perfusion,[2,3] which can be indicative of cancer hallmarks such as local hypoxia and angiogenesis.[4,5]

PAI is able to measure optical absorption by using pulsed laser light to induce the photoacoustic effect. The absorbed light creates an initial pressure distribution which gives rise to a sound wave. PA raw sensor data of this wave can be measured with a common US transducer.[6] To be able to access the information contained in the raw data, the initial pressure distribution needs to be reconstructed. While standard reconstruction algorithms such as the radon transform can be used for tomographic scans,[7] they can currently only be applied to small animal imaging,[8] but not to clinically used scanners[9] without the introduction of imaging artifacts. This is, because clinical scanners used for PA usually have a linear sensor geometry, therefore PA images have to be reconstructed from a partial scan of the emitted sound waves. This so-called limited view problem causes undersampling during image reconstruction, giving rise to artifacts and noise amplification.[10]

Further author information: (Send correspondence to J.G. or L.M.H.)
J.G.: E-mail: j.groehl@dkfz-heidelberg.de
L.M.H.: E-mail: l.maier-hein@dkfz-heidelberg.de

Photons Plus Ultrasound: Imaging and Sensing 2018, edited by Alexander A. Oraevsky, Lihong V. Wang, Proc. of SPIE Vol. 10494, 104942S · © 2018 SPIE · CCC code: 1605-7422/18/$18 · doi: 10.1117/12.2288353

Two state-of-the-art methods are currently mainly used for reconstructing PA images from raw sensor data. These are time reversal[11] and delay-and-sum (DAS)[12] beamforming. Time reversal is a method where a numerical model of the forward problem is simulated backwards in time and delivers approximations of the initial pressure distribution.[11] DAS beamforming was originally developed for US image reconstruction and is commonly used in PA. Its purpose is to yield the origins of the measured soundwaves.[13] It is a simple and fast algorithm but lacks the ability to provide quantitative estimations of the initial pressure distribution.[14,15]

While the vast majority of reconstruction methods is model-based, first attempts have been made to apply deep learning approaches - which have yielded breakthrough success in a number of medical image processing challenges[16–19] - to the reconstruction of PA images. For example, Antholzer et al.[20] have proposed a deep learning based reconstuction of PA images from a sparse representation of raw sensor data, Hauptmann et al.[10] have used an iterative approach using a deep learning neural network to reconstruct initial pressure based on a vessel segmentation of lung CT images, and Reiter et al.[21] have used a convolutional network with feed-forward fully connected layers to identify locations of PA point sources. To our knowledge, however, deep learning has not yet successfully been applied to the reconstruction of the initial pressure distribution directly from limited view PA raw sensor data using linear US transducers.

In this paper, we address the challenge of reconstructing the initial pressure distribution from PA raw sensor data with convolutional neural networks (CNNs) based on a modification of the widely used U-Net deep neural network architecture.[22] We investigate two different variants of our method: (1) CNN-based post-processing of limited view DAS beamformed images and (2) direct CNN-based reconstruction of the initial pressure from PA raw sensor data as illustrated in Fig 1, without previous assumptions such as spherical wave propagation or speed of sound. According to first *in silico* experiments, qualitative and quantitative improvements compared to reconstruction with DAS beamforming can be achieved.

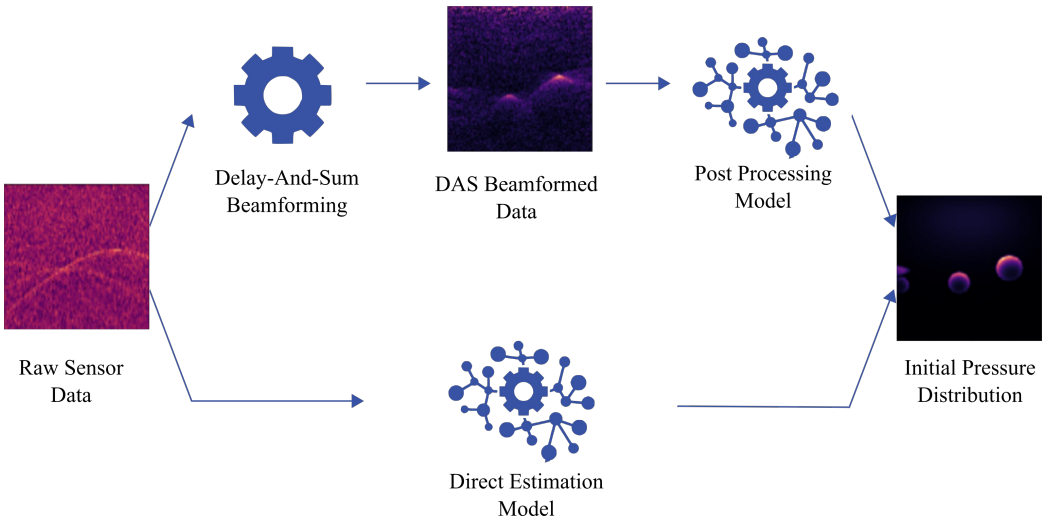


Figure 1. Visualization of two approaches to estimate initial pressure distributions from the PA raw sensor data. The first method (upper path) initially processes the PA raw sensor data with standard Delay-And-Sum (DAS) beamforming and the resulting image is post processed with a standard U-Net (post-processing model). The second approach (bottom path) can directly estimate the initial pressure distribution from the PA raw sensor data with a modified U-Net (direct estimation model).

2. METHODS

We tackle the acoustic inverse problem by estimating the initial pressure distribution from simulated PA raw sensor data. We investigate two different approaches using U-Net architectures (cf. Figure 1). In the first approach, we use a U-Net to post-process images reconstructed with DAS beamforming and in the second approach we estimate the initial pressure directly from the PA raw sensor data using a modified U-Net.

2.1 Image post-processing of DAS beamformed raw sensor data

We propose the use of a deep learning algorithm to estimate the initial pressure distribution from DAS reconstructions of PA raw sensor data. Specifically, we use a U-Net as it is the most successful and versatile state-of-the-art deep neural network[22] in medical image processing. The U-Net architecture is widely applied in image processing because of its ability to represent contextual and local image features on several levels of abstraction.[22] In our implementation the U-Net consists of three layers. The contracting steps consist of the repeated application of two convolutions with kernel size 3×3 and stride 1×1, each followed by a rectified linear unit (ReLu) and a batch normalization. The down-sampling is performed by a 2×2 max pooling operation. The expansive units consist of an upsampling followed by the same convolutions, ReLus and batch normalization steps as in the contracting units. Information from the contracting units is concatenated via the skip connections. For optimization we use PyTorch's[23] implementation of Adam with a base learning rate of 10^{-4} and a weight decay of 10^{-5} as well as a L1 loss function for back propagation.

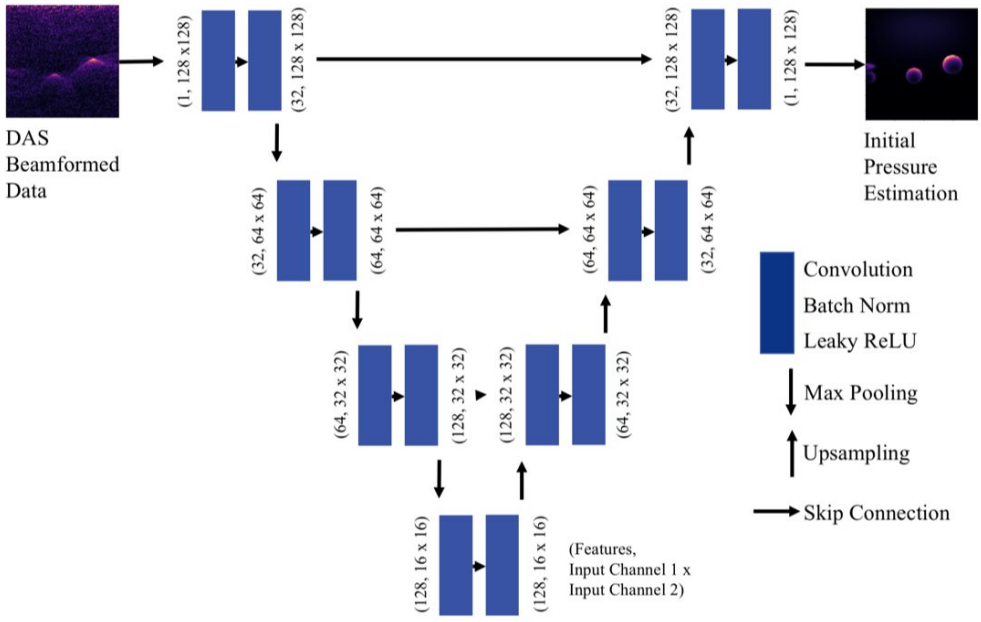


Figure 2. Architecture of the post-processing U-Net: The U-Net is trained to estimate the initial pressure distribution from the DAS beamformed images. Each computational unit, indicated by a blue box, consists of a convolution, a batch normalization and a leaky rectified linear unit. The downsampling is performed by max pooling operations and the upsampling is performed by upsampling operations. The number of feature maps is denoted as first number on the boxes, the image dimensions as second entry.

2.2 Direct estimation of the initial pressure from raw sensor data

A direct reconstruction of PA raw sensor data was performed with a modified U-Net architecture. In this case, we did not preprocess the PA raw sensor data with DAS beamforming but used the sensor data directly. The major challenge was to transform the data from a space and a time domain to a 2 dimensional space domain. Therefore, the skip connections of the U-Net shown in Figure 2 were altered by introducing a layer in each skip connection which reduces the image dimensions in one direction by using an anisotropic kernel size of 20×3 and striding of 20×1. The modified U-Net maps the information from the time axis to a spatial dimension and thereby reduces the resolution of the input PA raw sensor data from 2560×128 pixels to the target output resolution of 128×128 pixels.

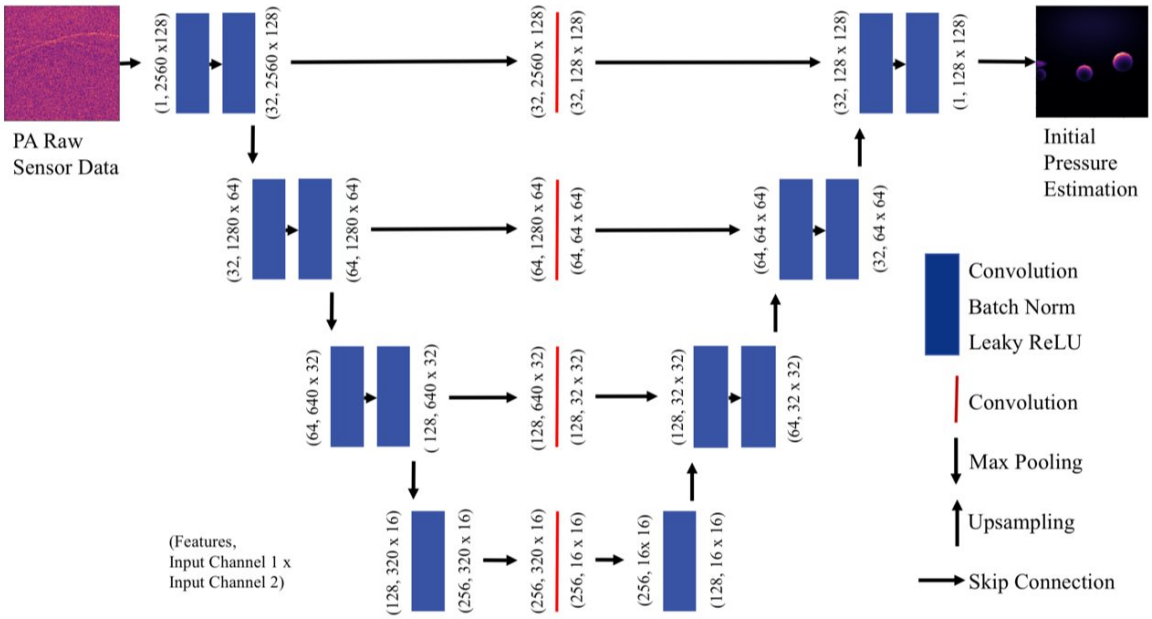


Figure 3. Architecture of the direct reconstruction U-Net. The modified U-Net is trained to estimate the initial pressure distribution directly from the PA raw sensor data. The skip connections of the U-Net include a convolutional layer that resizes the information to the target resolution. The contracting and expansive layers are the same as used with the standard U-Net.

2.3 Simulation of training data

In silico images of initial pressure distributions were simulated given a distribution of optical parameters using an adaptation of the widely used Monte Carlo simulation framework mcxyz.[24] The simulation pipeline is integrated in the Medical Imaging Interaction Toolkit (MITK)[25] and the process is explained in detail in a previous publication.[26]

Property	Value range	Unit
Reduced Scattering	15	cm^{-1}
Background Absorption	0.1	cm^{-1}
Vessel Absorption	U(2-8)	cm^{-1}
Vessel Radius	U(0.5 - 4)	cm
Number of Vessels	U(1 - 7)	

Table 1. Table with the tissue properties used in the mcxyz Monte Carlo Simulation. Here, U denotes a uniform distribution over the given value range.

To simulate raw sensor data from initial pressure distribution, we used the k-Wave framework with the k-Wave 2D FFT pseudo-spectral domain method for a line sensor. We used a virtual US probe with 128 transducer elements. To reduce artifacts produced by image borders we embedded our input images into a larger homogeneous pressure distribution. Our input images had a resolution of 128 x 128 pixels with an isotropic pixel spacing of 0.3 mm. The simulation depth for the PA raw sensor data was 57.6mm corresponding to 2560 Pixels with a time spacing of 1.5×10^{-8} s. The medium tissue density was approximated with 1000 kg/m^3 and the speed of sound was set to 1500 m/s.

3. EXPERIMENTS

We used a data set simulated as described above for training and validation, and a second dataset for testing of the two presented approaches. Using mcxyz and k-Wave, we simulated a data set with 3600 images of which 20% were used as test set, the rest was split 80/20 in training and validation set. We added random Gaussian noise with an amplitude of 20% of the mean maximum pixel intensity over all simulated image slices to the PA raw sensor data. The U-Net was implemented in python with the PyTorch framework[23] and we performed DAS beamforming with the MITK framework. Images in the training set were mirrored for data augmentation to increase the number training samples. Both models were trained with the same amount of data presented. The results reported in the following section were calculated on the test set while the validation set was used to supervise convergence of the training process and hyper-parameter optimization.

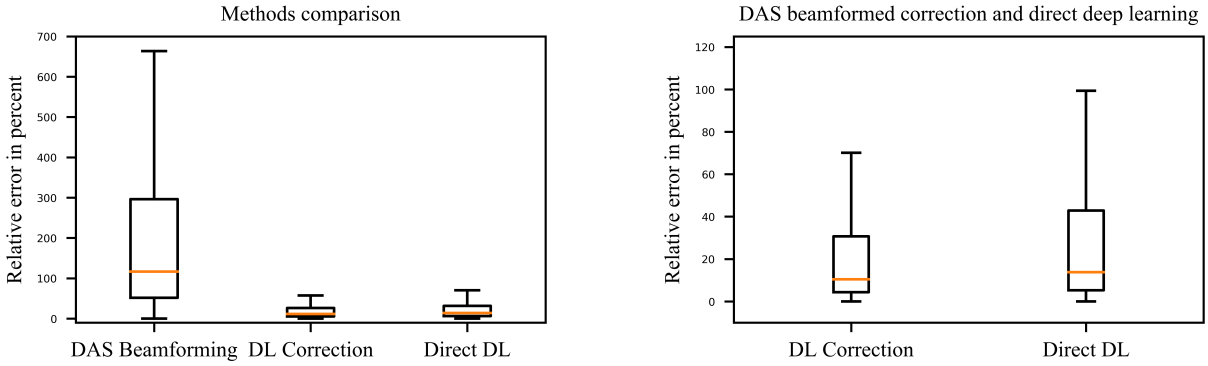
For the following results we calculated the relative initial pressure estimation error to the ground truth for each estimated image.

$$err = \left| \frac{\hat{IP} - IP}{IP} \right| \tag{1}$$

where err is the calculated error, $\hat{IP}$ is the initial pressure estimation and IP is the simulated ground truth initial pressure.

4. RESULTS

The quantitative comparison of the pixelwise mean difference on the entirety of the images is shown in Figure 4. The images have been normalized using z-score normalization over the whole dataset before the evaluation. The median relative initial pressure estimation error improved from 98% with an interquartile range (IQR) of [52% - 296%] in case of the DAS beamforming to 10% with IQR [0.5% - 26%] using the proposed deep learning architecture for correcting DAS beamformed images. The relative error of the direct deep learning approach was 14% with IQR [0.6% - 32%]. In a direct comparison, without prior image normalization, the median relative initial pressure estimation error was 12% with IQR [0.4% - 31%] in case of the correction of DAS beamformed images and the direct approach yields a median error of 14% with IQR [0.5% - 43%].

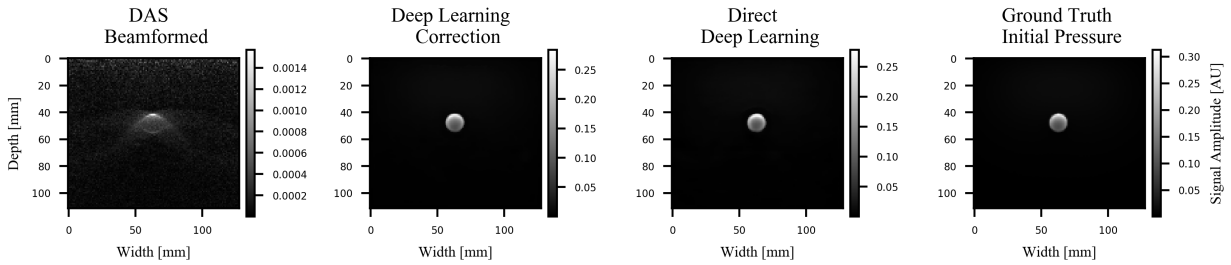


(a) Normalized images using z-score normalization for comparison of the suggested methods with DAS beamforming.

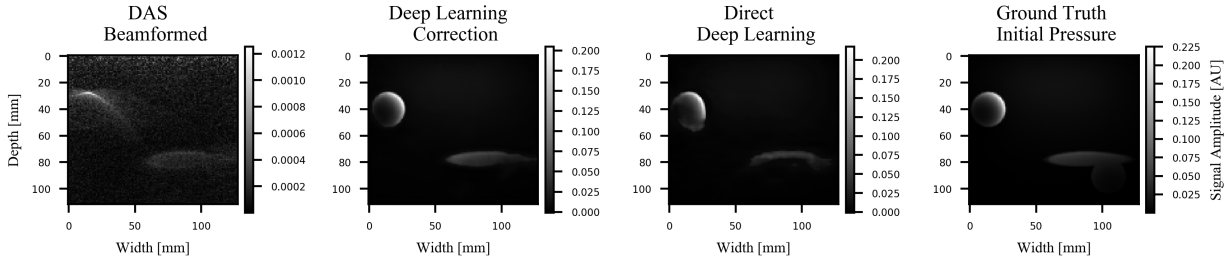
(b) Unnormalized images for direct comparison of the two methods.

Figure 4. Boxplots of the pixel-wise relative initial pressure distribution estimation errors. Both methods reveal advantages compared to DAS beamforming, as both deep learning based approaches achieved a median relative initial pressure estimation error of less than 15%.

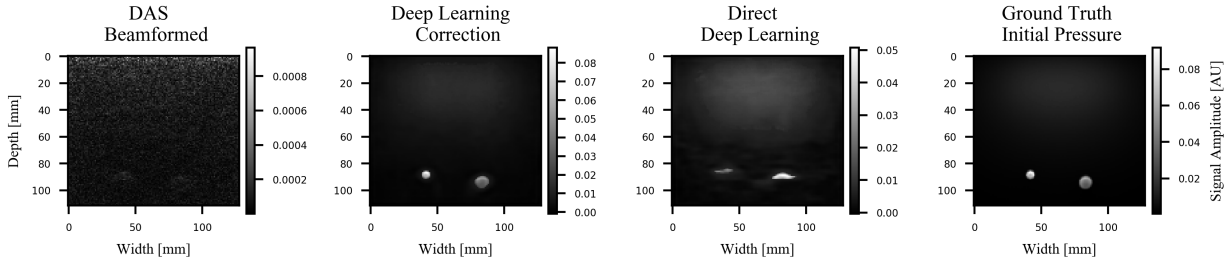
Figure 5 shows qualitative comparisons of DAS Beamformend images, of post processed images, directly reconstructed images and the ground truth simulated initial pressure distribution. To evaluate the best and worst performances we sorted the results on the test set by the distance to the ground truth. Only images containing vessel structures were taken into account.



(a) Image with the smallest error *err*. The signal is very close to the tissue surface and the shape is preserved well.



(b) Image with a median error *err*. Even though both different shaped vessels were reconstructed, the direct approach yields some inaccuracies especially in the lower parts.



(c) Image with the largest error *err*. The signal is very deep inside the tissue hardly visible in the DAS beamformed image and with large errors in the direct approach.

Figure 5. Sample images from the test set, where the directly reconstructed image from had the best, worst or median error *err*.

5. DISCUSSION

In this paper we present two methods that could improve currently used PAI reconstruction algorithms for limited view PA data. The first method was used as post-processing after Delay-And-Sum (DAS) beamforming of PA raw sensor data as also done by previous work.[10,21] We also demonstrate that in our dataset the DAS beamformed data can be corrected with state-of-the-art deep learning methods. In fact, the presented post-processing method achieved qualitative, as well as quantitative, improvements compared to DAS beamforming and yielded accurate results even under application of noise. However, one has to acknowledge that comparing DAS beamforming with the proposed methods is not fair, as it was not originally developed to yield a quantitative estimate of the initial pressure distribution.

Furthermore we showed that initial pressure distributions can be directly reconstructed from the PA raw sensor data with a modified U-Net without the use of any prior processing steps. This method needs one less calculation step but gives higher relative errors than the deep learning correction method. The main challenge was to transform the image dimensions from the PA raw sensor data to the initial pressure distribution. This was solved by introducing additional convolutional layer in the skip connections of the U-Net. As skip connections

were originally introduced to retain the entire range of high frequency information, further work could be done to improve the use of a U-Net for direct image reconstruction.

We have not yet tested the proposed methods on *in vivo* or *in vitro* data. This is because our *in silico* model does not include a realistic modelling of noise, angular sensitivity of the US transducer, the presence of artifacts and local variations of speed of sound and acoustic attenuation leading to a generally lower contrast to noise ratio. Due to the manner in which we trained and tested on random data from the same Monte Carlo distribution we cannot ensure that there exists no bias in the dataset. This fact should also be taken into account when comparing the results or our two algorithms to the DAS baseline approach. As such, our next challenge is to generalize these methods to be applicable to realistic data.

In limited view PA imaging it is essential to develop accurate and fast reconstruction methods. In this study we show two methods that can be used to either enhance reconstruction of DAS beamforming with a post-processing step or provide a direct reconstruction of the initial pressure distribution from PA raw sensor data. Results show high accuracy even with application of noise, while providing a qualitative improvement compared to beamforming. Future work will include comparing our methods to other advanced reconstruction algorithms such as done by Hauptmann et al.[10] and we will apply the methods to *in vitro* and *in vivo* data.

ACKNOWLEDGEMENTS

The authors would like to acknowledge support from the European Union through the ERC starting grant COMBIOSCOPY under the New Horizon Framework Programme grant agreement ERC-2015-StG-37960. The authors would also like to thank the ITCF of the DKFZ for the provision of their computing cluster.

REFERENCES

[1] Wang, L. V. and Yao, J., "A Practical Guide to Photoacoustic Tomography in the Life Sciences," *Nature methods* **13**, 627–638 (July 2016).

[2] Wang, L. V. and Gao, L., "Photoacoustic Microscopy and Computed Tomography: From Bench to Bedside," *Annual Review of Biomedical Engineering* **16**(1), 155–185 (2014).

[3] Pilatou, M. C., Kolkman, R. G. M., Hondebrink, E., Bolt, R. A., and Mul, F. F. M. d., "Photoacoustic imaging of blood perfusion in tissue and phantoms," **4256**, 28–34, International Society for Optics and Photonics (June 2001).

[4] Carmeliet, P. and Jain, R. K., "Angiogenesis in cancer and other diseases," *Nature* **407**, 35025220 (Sept. 2000).

[5] Zackrisson, S., Ven, S. M. W. Y. v. d., and Gambhir, S. S., "Light In and Sound Out: Emerging Translational Strategies for Photoacoustic Imaging," *Cancer Research* **74**, 979–1004 (Feb. 2014).

[6] Wang, K. and Anastasio, M. A., "Photoacoustic and Thermoacoustic Tomography: Image Formation Principles," in [*Handbook of Mathematical Methods in Imaging*], 781–815, Springer, New York, NY (2011). DOI: 10.1007/978-0-387-92920-0_18.

[7] Kunyansky, L. A., "Explicit inversion formulae for the spherical mean Radon transform," *Inverse Problems* **23**(1), 373 (2007).

[8] Xia, J., Yao, J., and Wang, L. V., "Photoacoustic tomography: principles and advances," *Electromagnetic waves (Cambridge, Mass.)* **147**, 1–22 (2014).

[9] Razansky, D., Distel, M., Vinegoni, C., Ma, R., Perrimon, N., Koester, R. W., and Ntziachristos, V., "Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo," *Nature Photonics* **3**, 412–417 (July 2009).

[10] Hauptmann, A., Lucka, F., Betcke, M., Huynh, N., Cox, B., Beard, P., Ourselin, S., and Arridge, S., "Model based learning for accelerated limited view 3d photoacoustic tomography," *arXiv:1708.09832 [cs, math]* (Aug. 2017). arXiv: 1708.09832.

[11] Treeby, B. E. and Cox, B. T., "k-Wave MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of Biomedical Optics* **15**, 021314 (Apr. 2010).

[12] Thomenius, K. E., "Evolution of ultrasound beamformers," in [*1996 IEEE Ultrasonics Symposium. Proceedings*], **2**, 1615–1622 vol.2 (Nov. 1996).

[13] Mozaffarzadeh, M., Mahloojifar, A., and Orooji, M., "Medical photoacoustic beamforming using minimum variance-based delay multiply and sum," **10335**, 1033522, International Society for Optics and Photonics (June 2017).

[14] Matrone, G., Savoia, A. S., Caliano, G., and Magenes, G., "The Delay Multiply and Sum Beamforming Algorithm in Ultrasound B-Mode Medical Imaging," *IEEE Transactions on Medical Imaging* **34**, 940–949 (Apr. 2015).

[15] Mozaffarzadeh, M., Mahloojifar, A., Orooji, M., Adabi, S., and Nasiriavanaki, M., "Double-Stage Delay Multiply and Sum Beamforming Algorithm: Application to Linear-Array Photoacoustic Imaging," *IEEE Transactions on Biomedical Engineering* **65**, 31–42 (Jan. 2018).

[16] Shen, D., Wu, G., and Suk, H.-I., "Deep Learning in Medical Image Analysis," *Annual Review of Biomedical Engineering* **19**(1), 221–248 (2017).

[17] Jäger, P. F., Bickelhaupt, S., Laun, F. B., Lederer, W., Heidi, D., Kuder, T. A., Paech, D., Bonekamp, D., Radbruch, A., Delorme, S., Schlemmer, H.-P., Steudle, F., and Maier-Hein, K. H., "Revealing Hidden Potentials of the q-Space Signal in Breast Cancer," in [*Medical Image Computing and Computer Assisted Intervention MICCAI 2017*], *Lecture Notes in Computer Science*, 664–671, Springer, Cham (Sept. 2017).

[18] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., and Snchez, C. I., "A Survey on Deep Learning in Medical Image Analysis," *Medical Image Analysis* **42**, 60–88 (Dec. 2017). arXiv: 1702.05747.

[19] Wirkert, S. J., Vemuri, A. S., Kenngott, H. G., Moccia, S., Götz, M., Mayer, B. F. B., Maier-Hein, K. H., Elson, D. S., and Maier-Hein, L., "Physiological Parameter Estimation from Multispectral Images Unleashed," in [*Medical Image Computing and Computer-Assisted Intervention MICCAI 2017*], *Lecture Notes in Computer Science*, 134–141, Springer, Cham (Sept. 2017).

[20] Antholzer, S., Haltmeier, M., and Schwab, J., "Deep learning for photoacoustic tomography from sparse data," *arXiv preprint arXiv:1704.04587* (2017).

[21] Reiter, A. and Bell, M. A. L., "A machine learning approach to identifying point source locations in photoacoustic data," **10064**, 100643J, International Society for Optics and Photonics (Mar. 2017).

[22] Ronneberger, O., Fischer, P., and Brox, T., "UNet Convolutional Networks for Biomedical Image Segmentation," *arXiv:1505.04597 [cs]* (May 2015). arXiv: 1505.04597.

[23] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A., "Automatic differentiation in PyTorch," (Oct. 2017).

[24] Jacques, S. L., "Coupling 3d Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation," *Photoacoustics* **2**, 137–142 (Dec. 2014).

[25] Wolf, I., Vetter, M., Wegner, I., Böttger, T., Nolden, M., Schöbinger, M., Hastenteufel, M., Kunert, T., and Meinzer, H.-P., "The Medical Imaging Interaction Toolkit," *Medical Image Analysis* **9**, 594–604 (Dec. 2005).

[26] Kirchner, T., Gröhl, J., and Maier-Hein, L., "Local context encoding enables machine learning-based quantitative photoacoustics," *arXiv:1706.03595 [physics]* (June 2017). arXiv: 1706.03595.