Deep Learning Image Reconstruction Method for Limited-Angle Ultrasound Tomography in Prostate Cancer

Alexis Cheng[1], Younsu Kim[2], Emran M.A. Anas[2], Arman Rahmim[2,3], Emad M. Boctor[2], Reza Seifabadi[1], Bradford J. Wood[1]

[1]National Institutes of Health, Bethesda, MD, USA
[2]John Hopkins University, Baltimore, MD, USA
[3]University of British Columbia, Vancouver, BC, Canada

ABSTRACT

Problem: The gold standard for prostate cancer diagnosis is B-mode transrectal ultrasound-guided systematic core needle biopsy. However, cancer is indistinguishable under ultrasound and thus additional costly imaging methods are necessary to perform targeted biopsies. Speed of sound is a potential biomarker for prostate cancer and has the potential to be measured using ultrasound tomography. Given the physical constraints of the prostate's anatomy, this work explores a simulation study using deep learning for limited-angle ultrasound tomography to reconstruct speed of sound.

Methods: A deep learning-based image reconstruction framework is used to address the limited-angle ultrasound tomography problem. The training data is generated using the k-wave acoustic simulation package. The general network structure is composed of a series of dense fully-connected layers followed by an encoder and a decoder network. The basic idea behind this neural network is to encode a time of flight map into a lower dimension representation that can then be decoded into a speed of sound image.

Results and Conclusions: We show that limited-angle UST is feasible in simulation using an auto-encoder-like DL framework. There was a mean absolute error of 7.5 ± 8.1 m/s with a maximum absolute error of 139.3 m/s. Future validation on experimental data will further assess their ability in improving limited-angle ultrasound tomography.

INTRODUCTION

Prostate cancer (PCa) is the second leading cause of cancer death in American men [1]. The gold standard for PCa diagnosis is B-mode transrectal ultrasound-guided systematic core needle biopsy. This procedure is blind however because tumors are generally indistinguishable from healthy prostate tissue in ultrasound (US) images. As a result, there are many different biopsy schemes that aim to best cover the entire prostate. These schemes are random in nature and are susceptible to undersampling the prostate, leading to less effective cancer detection [4]. It has also led to the development and use of targeted prostate biopsy methods using multiparametric magnetic resonance imaging (mpMRI) [2, 3] or US/mpMRI fusion [4, 5]. In these methods, suspicious tumor regions are diagnosed by a radiologist in the pre-operative mpMRI and are used during the procedure as biopsy targets.

Targeted prostate biopsies have the advantage in that the biopsies are now done with prior knowledge of suspicious regions of the prostate obtained from pre-operative mpMRI. Siddiqui et al. showed increased detection of high-risk prostate cancer and decreased detection of low-risk prostate cancer when using mpMRI/US fusion biopsy compared with standard extended-sextant ultrasound-guided biopsy [4]. However, given the assortment of high costs associated with using MRI, an alternative cost-effective imaging solution may be necessary for cases in which MRI is cost-prohibitive.

Medical Imaging 2019: Ultrasonic Imaging and Tomography, edited by Brett C. Byram, Nicole V. Ruiter, Proc. of SPIE Vol. 10955, 1095516 · © 2019 SPIE
CCC code: 1605-7422/19/$18 · doi: 10.1117/12.2512533

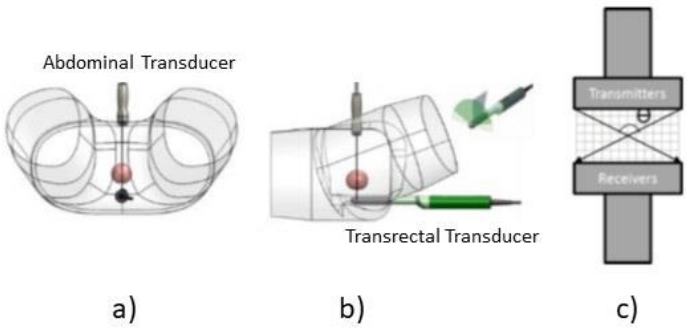

Figure 1. Limited-angle ultrasound tomography transducer scenarios a) abdominal probe and transrectal linear probe b) abdominal probe and transrectal curvilinear probe. c) 2 abdominal probes

Ultrasound tomography (UST) may be one such imaging solution. Ultrasound tomography is a quantitative imaging modality and has been most commonly used in breast imaging for the detection and diagnosis of breast cancer [6, 7]. It has been shown that UST can provide biomarkers, such as speed-of-sound (SoS) that can be used to characterize tissue types in the breast and differentiate between them [8]. UST systems designed for the breast typically resemble a ring or a hemisphere and cannot be used for prostate imaging without modification. This type of system represents a full-angle UST scenario where the acoustic waves travel across the region or organ of interest from approximately all directions. We previously demonstrated a robotics-based UST system that may be generalizable to any organ [9]. In this system shown in figure 1, two US transducers face each other in a limited-angle scenario as the acoustic waves travel across the region of interest from a limited number of directions. Figure 1a and 1b show a two-transducer setup that may be applicable to prostate imaging, combining an abdominal US transducer with an endorectal US transducer. Figure 1c shows the initial experimental setup using two abdominal transducers to emulate the prostate scenario. Thus far, many UST image reconstruction methods have focused on full-wave inverse scattering [10, 11], which can be difficult to solve efficiently and accurately as the limited-angle scenario is an ill-posed condition problem.

There have been many recent developments in the field of deep imaging. In the context of medical imaging, until recently, deep learning (DL) and machine learning (ML) have been used mainly for labeling and classification tasks [12-14]. Deep imaging is the use of DL to learn the underlying model of the imaging sensor, allowing one to recover the image from the sensor data. There have been some promising results for deep imaging in MRI and CT sparse data reconstruction problems [15-17]. Inspired by these promising results and the problems with using full-wave inverse scattering in this limited-angle scenario, we aim to explore a DL-based framework to solve the limited-angle UST problem.

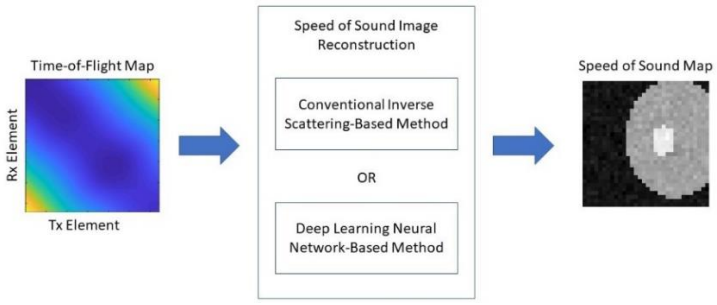

Figure 2. Speed of sound image reconstruction workflow: conventional inverse scattering-based method or deep learning neural network-based method

Figure 2 shows how we propose the DL-based framework to be integrated into the image reconstruction process. The core of image reconstruction is recovering the image from the sensor data. As such, conventional image reconstruction uses inversion methods based on physical knowledge of the sensor. We propose to use a DL-based framework in place of these inversion methods.

This manuscript will describe a DL-based framework that takes in time-of-flight (ToF) measurements derived from the sensor data as the input and returns a SoS image of the medium as the output. The methods section will describe the DL network structure and how one can train it. The results section will show simulation results and an analysis on the resulting errors. Following that, we will discuss ways that such a DL-based limited-angle UST image reconstruction framework may be improved.

1. METHODS

1.1 Deep Learning Training

One important aspect of training deep learning networks is the presence of training data. Conventionally, DL networks are trained with labeled data, where one has knowledge of the output corresponding with each set of collected input data. These labels are often created manually. For example, in a segmentation task, one might ask experts to segment or label many images to train a DL network.

In the context of image reconstruction, knowledge of the medium or region being imaged must be obtained. For UST, this corresponds to the SoS image of the medium. In general, it is very difficult to know the SoS image of the medium without dedicated devices for determining the SoS. One option is the use of phantoms with known and precise SoS parameters, but it is impractical to create thousands of phantoms for the purposes of collecting training data.

Simulation is one solution to this problem. The requirement for using simulation to generate training data is that one must have knowledge of the forward model to the system. In UST, this means that given a SoS image of the medium, we can determine the data received by the sensors which corresponds to knowledge of the sensor locations as well as the propagation of the acoustic waves. We randomly generate SoS images based on the parameters shown in table 1. These include speed of sounds for a background, a prostate, and 2 regions of interest (ROI – Fig. 3). The speed of sound values of each region is composed of 2 components. The first component is a baseline speed of sound drawn from a uniform distribution. The second component is a value drawn from a normal distribution to simulate pixel inhomogeneity. These values are chosen based on our assumption on the expected SoS in prostate tissue drawn from literature and a previous study [18, 19]. The prostate and ROI shapes are simulated as ellipses based on equation 1 as a representation of what one might see in the body without using complicated higher order shapes. Higher order shapes are also possible but were not investigated at this time. $r_{p,c}$ and $e_{p,c}$ represent the radius and ellipsoidal factors of the prostate and ROIs respectively. x_p, and y_p represent the center of the prostate. An equation replacing the prostate parameters with the ROI parameters can be easily derived from equation 1. We also note here that the DL network must be retrained with the corresponding training data for different sensor or imaging configurations. A few examples of the simulated SoS images can be seen in figure 3.

Table 1. Speed of sound simulation parameters and values

$$\frac{(x-x_p)^2}{e_{px}^2} + \frac{(y-y_p)^2}{e_{py}^2} \leq r_p^2 \qquad (1)$$

Parameter	Value
Background speed of sound	U[1490,1540] + 5*N(0,1)
Prostate speed of sound (SoS_p)	U[1515,1615] + 5*N(0,1)
Prostate radius (r_p)*	U[15, 30]
Prostate ellipsoidal factor (e_p)*	U[0.9, 1]
Region of interest speed of sound (2)	SoS_p+50*U[-0.5,0.5]+5*N(0,1)
ROI radius (r_c)*	U[1, 2]
ROI ellipsoidal factor (e_c)*	U[1, 4]

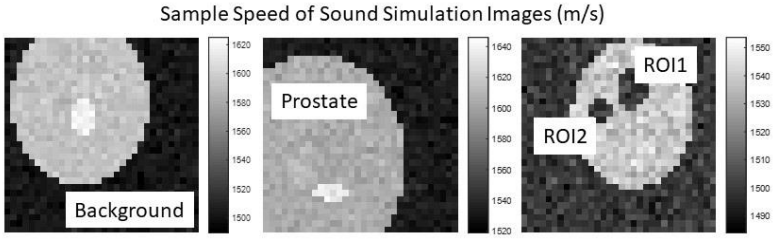

Figure 3. Speed of Sound Simulation Examples for training the network

This SoS image is then used alongside a simulated acoustic environment representing the experimental platform to generate the corresponding ToF image. The acoustic environment consists of two 6-cm 32-element transducers facing each other separated by 5 cm. We do this process with the MATLAB k-Wave toolbox [20]. A total of 5000 images were used, with half for training and half for validation.

1.2 Network Structure

Encoder Network

Layer (type)	Output Shape	Param #
encoder_input (InputLayer)	(None, 32, 32, 1)	0
flatten_1 (Flatten)	(None, 1024)	0
dropout_1 (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 1024)	1049600
activation_1 (Activation)	(None, 1024)	0
dropout_2 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 1024)	1049600
activation_2 (Activation)	(None, 1024)	0
reshape_1 (Reshape)	(None, 32, 32, 1)	0
conv2d_1 (Conv2D)	(None, 16, 16, 64)	1664
batch_normalization_1 (Batch	(None, 16, 16, 64)	256
conv2d_2 (Conv2D)	(None, 8, 8, 64)	102464
batch_normalization_2 (Batch	(None, 8, 8, 64)	256
conv2d_3 (Conv2D)	(None, 4, 4, 64)	102464
batch_normalization_3 (Batch	(None, 4, 4, 64)	256
flatten_2 (Flatten)	(None, 1024)	0
latent_vector (Dense)	(None, 16)	16400

Total params: 2,322,960
Trainable params: 2,322,576
Non-trainable params: 384

Decoder Network

Layer (type)	Output Shape	Param #
decoder_input (InputLayer)	(None, 16)	0
dense_3 (Dense)	(None, 1024)	17408
latent_activation (Activatio	(None, 1024)	0
reshape_2 (Reshape)	(None, 4, 4, 64)	0
conv2d_transpose_1 (Conv2DTr	(None, 8, 8, 64)	102464
batch_normalization_4 (Batch	(None, 8, 8, 64)	256
conv2d_transpose_2 (Conv2DTr	(None, 16, 16, 64)	102464
batch_normalization_5 (Batch	(None, 16, 16, 64)	256
conv2d_transpose_3 (Conv2DTr	(None, 32, 32, 64)	102464
batch_normalization_6 (Batch	(None, 32, 32, 64)	256
conv2d_transpose_4 (Conv2DTr	(None, 32, 32, 1)	1601

Total params: 327,169
Trainable params: 326,785
Non-trainable params: 384

Complete Network

Layer (type)	Output Shape	Param #
encoder_input (InputLayer)	(None, 32, 32, 1)	0
encoder (Model)	(None, 16)	2322960
decoder (Model)	(None, 32, 32, 1)	327169

Total params: 2,650,129
Trainable params: 2,649,361
Non-trainable params: 768

Figure 4. Deep learning network structure description: Various layers and number of trainable parameters in encoder, decoder, and complete network respectively

Figure 4 shows the network structure that we used in this study. The structure resembles an auto-encoder [21]. The main function of an auto-encoder is to compress or encode the higher dimension input data into a lower dimension parameter vector that represents the input data and determine the most significant parameters that define the input. To accomplish this, an auto-encoder will have an encoder half and a decoder half. As the names suggest, one aims to encode the higher dimension input data into a lower dimension parameter vector while the other aims to decode the lower dimension parameter vector into a higher dimension output data.

The use of an auto-encoder in UST image reconstruction is slightly different than its conventional use for image denoising where the input and output data are in the same domain or image space. In this scenario, the input and output data to the auto-encoder DL network are no longer the same as the input is the ToF image whereas the output is the SoS image. Instead, the aim is to encode the input sensor ToF data into a lower dimension parameter vector, that can then be decoded into the output SoS image. The basic idea of this network is like AUTOMAP by Zhu et al. [16], which was used for MRI and PET image reconstruction. We implemented this network in Keras [22], with underlying support by tensorflow [23].

2. RESULTS

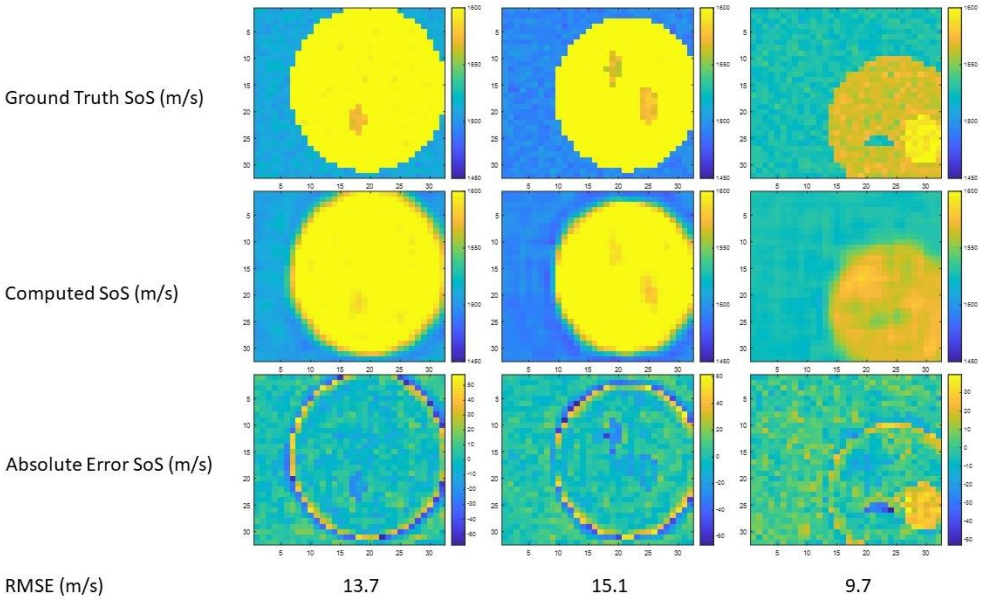

Figure 5. Sample validation images comparison between ground truth and computed speed of sound and the absolute error between them

Figure 5 shows a set of validation SoS images computed with the DL-based UST framework compared with the ground truth simulated SoS map. Quantitatively, there was an absolute error of 7.5 ± 8.1 m/s with a maximum absolute error of 139.3 m/s. The errors are particularly evident at the borders of regions with differences in speed of sound.

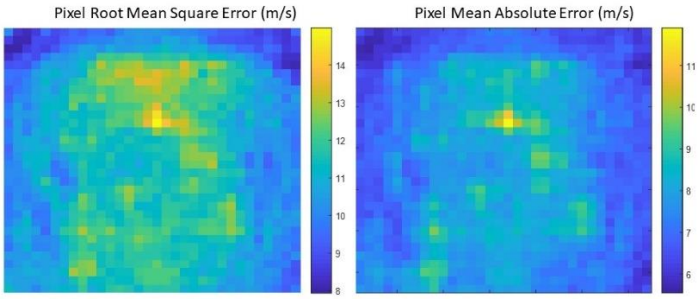

Figure 6. Pixel-wise root mean square errors and mean absolute error within imaging region

Figure 6 shows the pixel-wise root mean square error and mean absolute error at each spatial location relative to the imaging apparatus across all 2500 of the validation images.

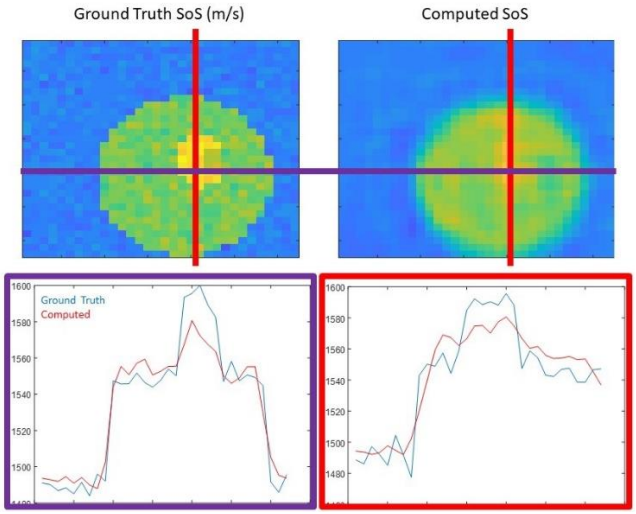

Figure 7. Comparison of resolution on a sample validation image in axial and lateral dimensions

Figure 7 shows the line plots across both dimensions of a sample validation SoS image reconstructed with the DL-based UST framework compared with the ground truth simulated SoS map.

3. DISCUSSION

From the results in figure 5, we can see that a DL-based limited-angle UST reconstruction method is feasible with certain limitations. First, from the error difference images, it is observed that while the main SoS regions are reconstructed well with mean errors under 10 m/s, most of the errors are present at the transitions between different speed of sound bodies. This is partially attributed to a loss in resolution in the computed images as can be seen in figure 7. The sharp edges in the original ground truth speed of sound images are smoothed out in the computed images. This causes a direct subtraction between the two images to have large discrepancies at these edges. An alternate way of comparison to explore in future work may be to first segment the computed images for the top discernible speed of sound regions and comparing that directly with the known speed of sound regions.

Figure 6 shows the error as a function of the image pixel's location relative to the imaging sensor. One would expect the central regions of the image to have lower errors because more acoustic paths pass through these

regions, leading to more knowledge of these regions. However, we can see that this is not the case in figure 6, and errors seem to be slightly higher in the central regions. The reason for this can also be explained by the loss in resolution described in the previous paragraph in conjunction with the prostate region being placed central in the image. The large errors observed at the region interfaces are concentrated in the central regions because that is where the region interfaces are also concentrated. A choice of more random speed of sound validation images may be one way of observing if the expectation that the central regions have lower speed of sound is valid.

The next step to validate this method is to test the trained DL-based image reconstruction framework on experimentally collected data. One of the main challenges is that any discrepancies in the imaging environment and the simulated environment used for training will result in errors. The DL-based image reconstruction framework is only trained to perform on the data used to train it. This challenge may lead to the need for many UST image reconstruction neural networks to be trained for each independent imaging scenario. One possible solution to decrease the amount of necessary training may be transfer learning, briefly training pre-trained networks with data representing the new imaging scenario.

4. CONCLUSION

This manuscript explores the use of a deep learning neural network applied to the limited-angle ultrasound tomography problem. We demonstrate that this DL-based image reconstruction method is feasible in simulation. Future work will extend this framework to experimentally collected data.

ACKNOWLEDGEMENTS

This work was supported by NIH Bench to Bed (B2B) Award. NIH owns intellectual property related to this topic. This work utilized the computational resources of the NIH HPC Biowulf cluster. (http://hpc.nih.gov).

REFERENCES

1. Torre L.A., Bray F., Siegel R.L., Ferlay J., Lortet-Tieulent J., and Jemal A., “Global cancer statistics, 2012”, in CA Cancer J. Clinical, 65(2), pp. 87-108, 2015.
2. Siddiqui M.M., Truong H., Rais-Bahrami S., Stamatakis L., Logan J., Walton-Diaz A., Turkbey B., Choyke P., Wood B.J., Simon R., and Pinto P.A., “Clinical implications of a multiparametric MRI based nomogram applied to prostate cancer active surveillance”, in Journal of Urology, 193(6), pp. 1943-1949, 2015.
3. Komai Y., Numao N., Yoshida S., Matsuoka Y., Nakanishi Y., Ishii C., Koga F., Saito K., Masuda H., Fujii Y., Kawakami S., and Kihara K., “High diagnostic ability of multiparametric magnetic resonance imaging to detect anterior prostate cancer missed by transrectal 12-core biopsy”, in Journal of Urology, 190(3), pp. 867-873, 2013.
4. Siddiqui M.M., Rais-Bahrami S., Turkbey B., George A.K., Rothwax J., Shakir N., Okoro C., Raskolnikov D., Parnes H.J., Linehan W.M., Merino M.J., Simon R.M., Choyke P.L., Wood B.J., and Pinto P.A., “Comparison of MR/ultrasound fusion-guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer”, in Journal of the American Medical Association, 313(4), pp. 390-397, 2015.
5. Lee D.J., Recabal P., Sjoberg D.D., Thong A., Lee J.K., Eastham J.A., Scardino P.T., Vargas H.A, Coleman J., and Ehdaie B., “Comparative Effectiveness of Targeted Prostate Biopsy Using MRI-US Fusion Software and Visual Targeting: a Prospective Study”, in Journal of Urology, 196(3), pp. 697-702, 2016.

6. Lenox M.W., Wiskin J., Lewis M.A., Darrouzet S., Borup D., and Hsieh S., “Imaging Performance of Quantitative Transmission Ultrasound”, in International Journal of Biomedical Imaging, vol. 2015, 454028, 2015.
7. Li C., Duric N., Littrup P., and Huang L., “In vivo breast sound-speed imaging with ultrasound tomography”, in Ultrasound in Medicine and Biology, 35(10), pp. 1615-1628, 2009.
8. Malik B., Klock J., Wiskin J., and Lenox M., “Objective breast tissue image classification using Quantitative Transmission ultrasound tomography”, in Scientific Reports, vol. 6, 38857, 2016.
9. Aalamifar F., Seifabadi R., Bernardo M., Negussie A.H., Turkbey B., Merino M., Pinto P., Rahmim A., Wood B.J., and Boctor E.M., “Ultrasound Tomosynthesis: A New Paradigm for Quantitative Imaging of the Prostate”, in Proceedings of International Conference on Medical Image Computing and Computer-Assisted Interventions, pp. 577-584, 2016.
10. Wiskin J., Borum D.T., Johnson S.A., and Berggren M., “Non-linear inverse scattering: high resolution quantitative breast tissue tomography”, in Journal of the Acoustical Society of America, 131(5), pp. 3802-3813, 2012.
11. Wang K., Matthews T., Anis F., Li C., Duric N., and Anastasio M.A., “Waveform Inversion with Source Encoding for Breast Sound Speed Reconstruction in Ultrasound Computed Tomography”, in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 62(3), pp. 475-493, 2015.
12. Zhang W., Li R., Deng H., Wang L., Lin W., Ji S., and Shen D., “Deep convolutional neural networks for multi-modality isointense infant brain image segmentation”, in Neuroimage, 108, pp. 214-224, 2015.
13. Shin H., Roberts K., Lu L., Demner-Fushman D., Yao J., and Summers R.M., “Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation”, in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
14. Shen D., Wu G., and Suk H., “Deep Learning in Medical Image Analysis”, in Annual Review of Biomedical Engineering, 19, pp. 221-248, 2017.
15. Wang G., “A Perspective on Deep Imaging”, in IEEE Access, 4, pp. 8914-8924, 2016.
16. Zhu B., Liu J.Z., Cauley S.F., Rosen B.R., and Rosen M.S., “Image reconstruction by domain-transform manifold learning”, in Nature, 555, pp. 487-497, 2018.
17. Zhang Z., Liang X., Dong X., Xie Y., and Cao G., “A Sparse-View CT Reconstruction Method Based on Combination of DenseNet and Deconvolution”, in IEEE Transactions on Medical Imaging, 37(6), pp. 1407-1417, 2018.
18. Tanoue H., Hagiwara Y., Kobayashi K., and Saijo Y., “Ultrasonic tissue characterization of prostate biopsy tissues by ultrasound speed microscope”, in Proceedings of IEEE Engineering in Medicine and Biology Society, pp. 8499-8502, 2011.
19. Seifabadi R., Cheng A., Malik B., Kishimoto S., Wiskin J., Munasinghe J., Negussie A.H., Bakhutashvili I., Cherukuri M., Choyke P., Pinto P., Rahmim A., Boctor E.M., Merino M., Lenox M., Turkbey B., and Wood B.J., “Correlation of ultrasound tomography with pathology and MRI in localizing prostate cancer”, presented at Ultrasonic Imaging and Tissue Characterization 2018.
20. Treeby B.E., and Cox B.T., “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields”, in Journal of Biomedical Optics, 15(2), 021314, 2010.
21. Liou C., Huang J., and Yang W., “Modeling word perception using the Elman network”, Neurocomputing, 71(16-18), 3150-3157, 2008.
22. Chollet F., and others, “Keras”, https://keras:io, 2015.
23. Abadi M. and others, “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems”, in Proceedings on Operating Systems Design and Implementation 2016., pp. 265-283. 2016.