A machine learning approach to identifying point source locations in photoacoustic data

Austin Reiter,[a] Muyinatu A. Lediju Bell[*b,c]

[a]Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA
[b]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA
[c]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

Interventional applications of photoacoustic imaging often require visualization of point-like targets, including the circular cross sectional tips of needles and catheters or the circular cross sectional views of small cylindrical implants such as brachytherapy seeds. When these point-like targets are imaged in the presence of highly echogenic structures, the resulting photoacoustic wave creates a reflection artifact that may appear as a true signal. We propose to use machine learning principles to identify these type of noise artifacts for removal. A convolutional neural network was trained to identify the location of individual point targets from pre-beamformed data simulated with k-Wave to contain various medium sound speeds (1440-1640 m/s), target locations (5-25 mm), and absorber sizes (1-5 mm). Based on 2,412 randomly selected test images, the mean axial and lateral point location errors were 0.28 mm and 0.37 mm, respectively, which can be regarded as the average imaging system resolution for our trained network. This trained network successfully identified the location of two point targets in a single image with mean axial and lateral errors of 2.6 mm and 2.1 mm, respectively. A true signal and a corresponding reflection artifact were then simulated. The same trained network identified the location of the artifact with mean axial and lateral errors of 2.1 mm and 3.0 mm, respectively. Identified artifacts may be rejected based on wavefront shape differences. These results demonstrate strong promise to identify point targets without requiring traditional geometry-based beamforming, leading to the eventual elimination of reflection artifacts from interventional images.

1. INTRODUCTION

Photoacoustic image quality is highly dependent on the quality of the beamforming process. Historically, beamformers have treated acoustic wave propagation as a geometry-based, time-of-flight measurement that accounts for signal time of arrival differences. However, these time-of-flight measurements do not consider that signals can be reflected multiple times and each acoustic reflection will be measured by the ultrasound transducer, thereby mapping reverberant signals to the wrong location. Two examples of interventional applications where acoustic reflections are a significant impediment to progress and clinical translation include the identification of brachytherapy seeds for treating prostate cancer[1,2] and the guidance of minimally invasive surgeries to avoid critical structures hidden by bone.[3,4]

Several methods are being developed to remove photoacoustic artifacts and enhance photoacoustic signals, including methods that incorporate singular value decomposition[5] and short-lag spatial coherence.[6] However, these approaches still rely on the same geometry-based, time-of-flight measurements as those of historical beamformers, and therefore similarly fail in the presence strong artifacts caused by bright, hyperechoic reflectors. A method named photoacoustic-guided focused ultrasound (PAFUSion)[7] was recently developed to address this challenge by using ultrasound data to mimic the wavefields produced by photoacoustic sources and thereby identify reflection artifacts for removal. This method depends on the assumption of identical acoustic reception pathways for ultrasound and photoacoustic data, which is not always true. In addition to the multiple paths that an acoustic signal can take once it encounters a strong reflector, there are multiple variables that govern acoustic wave propagation and the associated time-of-flight, such as tissue properties (e.g. sound speed, attenuation, absorption), tissue type (e.g. fat, bone, muscle), tissue thickness, and refraction.

*E-mail:mledijubell@jhu.edu

Photons Plus Ultrasound: Imaging and Sensing 2017, edited by Alexander A. Oraevsky, Lihong V. Wang, Proc. of SPIE Vol. 10064, 100643J · © 2017 SPIE · CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2255098

Deep learning, particularly deep learning with convolutional neural networks (CNNs), has in recent years redefined the state-of-the-art in computer vision and machine learning,[8–10] emerging as a quintessential technique to train models from a large database of possibilities, representing a characteristic that is ideally suited and uniquely applicable to the beamforming process. For example, although CNNs have only recently come to fruition, neural networks date back even further, and Nikoonahad et al.[11] used neural networks to remove artifacts in ultrasound images by estimating delay functions based on time-of-flight information. While this approach targets artifacts caused by sound speed estimation errors, it does not address the severely problematic artifacts caused by acoustic reflections.

We are developing deep learning methods to remove the reflection artifacts often seen in interventional photoacoustic imaging by eliminating receive beamforming and replacing it with a novel approach that uses CNNs. This approach is based on the observation that signals from point sources located at the same depth generally have similar wavefront shapes, regardless of the surrounding medium properties. The first step toward using machine learning to remove reflection artifacts is to accurately identify these point source locations. Toward this end, we trained a CNN to identify point source locations using simulated data. We then applied the trained network to an independent set of simulated data containing point sources and reflection artifacts, and additionally tested our trained CNN with real experimental data. The resulting point source locations may either be displayed separately or overlaid on co-registered ultrasound, CT, and/or MRI images, which are common imaging modalities for image-guided interventions and surgical navigation. The results may additionally be overlaid on photoacoustic images displayed with more traditional beamforming methods. To the authors' knowledge, this work represents the first application of machine learning to photoacoustic beamforming.

2. METHODS

We trained our CNN using simulated photoacoustic channel data and the widely available AlexNet network architecture.[8] The photoacoustic channel data was created with k-Wave simulations[12,13] and regarded as an input image. A total of 24,120 input images containing various point locations (axial range: 5-20 mm, lateral range: 5-25 mm, increment: 5 mm), various surrounding medium sound speeds (range: 1440-1640 m/s, increment: 1 m/s) causing known errors in the point location, and various point target radii (range: 1-5 mm, increment: 1 mm) were created in simulation. These images were randomly split into training (80%), validation (10%), and test (10%) sets. The output of the algorithm contained the point target 2D location in images created under these conditions, as illustrated in Fig. 1. The steps taken to achieve the 2D point location included nonlinearity with rectified linear units, max pooling, and regression with L2 normalization. The simulated photoacoustic images were divided into training and test sets. The trained network was applied to the test set of the simulated k-Wave photoacoustic data as well as to experimental data from a 3-mm diameter vessel-mimicking target obtained from a previous experiment using a pulsed laser diode.[14]

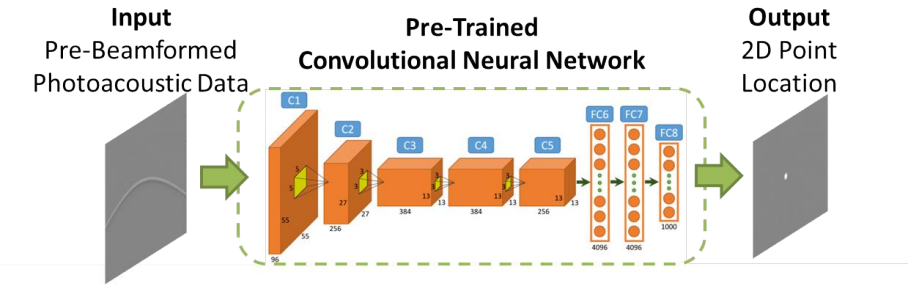

Figure 1: A convolutional neural network was trained to relate pre-beamformed acoustic data to point source locations.

3. RESULTS & DISCUSSION

3.1 One Simulated Source

The results of 2,412 test set images revealed that the average point location error was 0.28 mm and 0.37 mm in the axial and lateral dimensions, respectively. The mean errors generally decreased as axial and lateral distance increased, as illustrated with the box-whisker plots of Fig. 2. The top and bottom of each box represents the 75th and 25th percentiles of the measurements, respectively. The line inside each box represents the median measurement, and the whiskers (i.e., lines extending above and below each box) represent the range. Outliers were defined as any value greater than 1.5 times the interquartile range and are displayed as dots.

The errors reported in Fig. 2 can be regarded as the resolution of our imaging system when the known variation in sound speed ranges from 1440-1640 m/s and the known variation in target size ranges from 1-5 mm. With the traditional beamforming assumption of a fixed 1540 m/s sound speed, a medium known to consist of sound speeds within ±100 m/s of the assumed sound speed would cause the traditional images to have sound speed errors that equate to maximum distance errors of 6.5%. Thus, the maximum expected axial distance errors for this range of sound speeds is 0.3-1.3 mm (for the range of axial depths investigated), which is larger than the defined axial error with our imaging approach (Fig. 2(a)), particularly at the larger axial depths.

These results suggest that the proposed algorithm has excellent potential to detect points in the presence of clutter and other artifacts caused by speed of sound errors from multiple tissue layers. Once the network is trained, a point can be located at a frame rate of approximately 60 Hz when running this algorithm on our NVIDIA TITAN Black GPU.

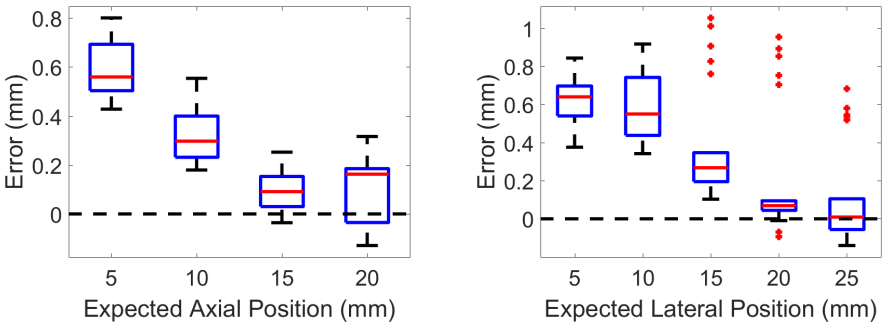

Figure 2: Summary of point location errors obtained with the trained CNN in the (a) axial and (b) lateral dimensions of the image.

3.2 Two Simulated Sources

To investigate the feasibility of detecting more than one point source in a single image, our trained CNN was applied to 20 images containing two signal sources. Our trained CNN successfully identified the source locations with mean axial and lateral errors of 2.6 mm and 2.1 mm, respectively. Because the network was only trained to detect single point sources, it is remarkable that we detected the location of more than one source in a single image. To achieve this detection, we noticed that the network was biased toward detecting the right-most wavefield. Thus, as demonstrated in Fig. 3 (top), the images were first input as normal (Input 1), then flipped from left to right, and the flipped image was treated as a new input (Input 2).

A true signal and a corresponding reflection artifact were then simulated in 20 images, as demonstrated in Fig. 3 (bottom). Note that the two wavefronts have different shapes, although they appear at the same depth, which indicates that one of these wavefronts originates from an artifact. The approach described above was used to identify the location of the simulated reflection artifact with mean axial and lateral errors of 2.1 mm and 3.0 mm, respectively. Although our trained CNN identified the locations of both the sources and the reflection artifacts with similar accuracy, it was unable to differentiate true signals from artifacts because it was not yet trained to detect these differences, although this is the end goal of the proposed technique.

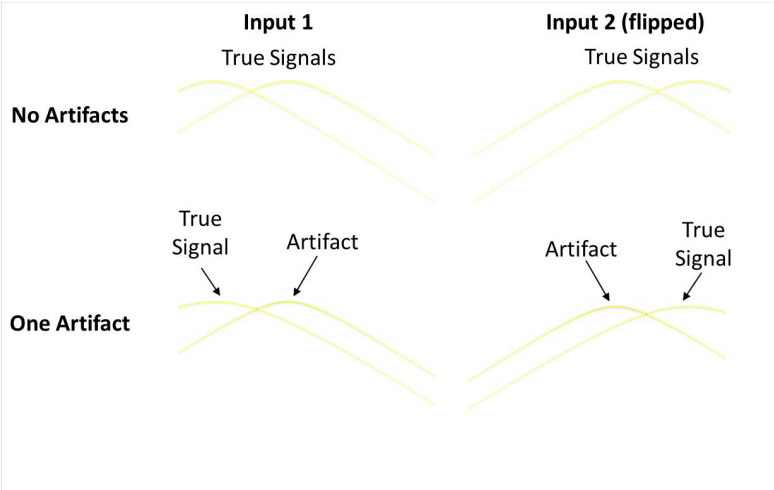

Figure 3: Examples of two point sources in a single image.

3.3 Artifact Detection

One possible approach to differentiate artifacts from true signals is to feed the identified locations of these two wavefields as inputs into an independent k-Wave simulation, then compare the resulting wavefields to those of the original image. Artifacts may then be rejected based on wavefront shape differences because our CNN will be trained to identify the locations of all true and reflected wavefields. This approach is reasonable because artifacts tend to manifest as obviously different wavefield shapes from true signals when they appear at an incorrect depth. One option to make this distinction is demonstrated in Fig. 4 where we analyze the shape of the waves at the peak location and construct histograms of the axial positions of the wavefield raw data (i.e. the image y-coordinates). Because our CNN is trained to predict the wavefield locations of both true signals and artifacts, we can reasonably train a classifier to distinguish between these two cases.

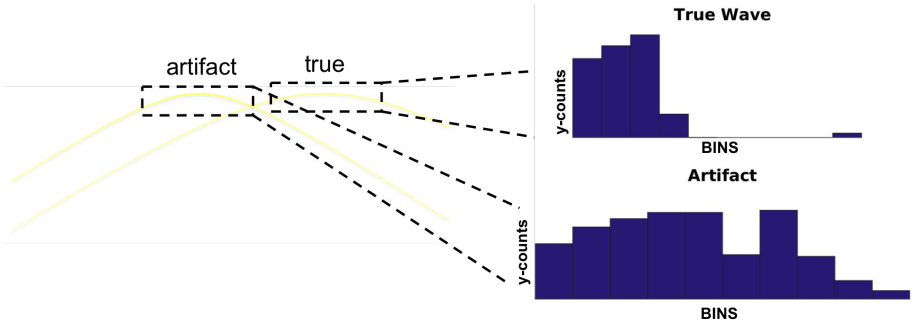

Figure 4: Demonstration of a histogram-based wavefield shape matching technique. The histogram shows the axial location of the wavefield segments in the dashed boxes. Note the differences when comparing the histogram of the true wave (top) to that of the artifact (bottom).

3.4 Application to Experimental Data

To demonstrate that our approach has broad applicability to multiple types of data, we tested our trained CNN on experimental data, which has differing characteristics when compared to simulated data. These differences can be appreciated by comparing Fig. 5 with Fig. 3. Note the noisier background, the incomplete wavefield arc, and the nonuniform signal intensities across the wavefield in Fig. 5, which shows the raw channel data from a 3-mm diameter rod submerged in water.[14] The axial and lateral positions of the rod with respect to the left corner of the ultrasound probe (which corresponds to the upper left corner of the photoacoustic image) was 6 mm and 11 mm, respectively. Our trained CNN estimated these positions within ∼2 mm in each dimension,

Figure 5: Experimental data from a 3 mm diameter blood vessel phantom.

identifying the point source locations as 4.21 mm (axial) and 13.01 mm (lateral). This estimation is within the errors obtained with this network, as reported in Section 3.2.

The error is larger than that obtained with the single source results (Fig. 2) likely due to the partiality of the wavefield, which has a similar appearance to one side of the full wavefield observed when two wavefields overlap (Fig. 3). Although additional training with partial wavefields might be required to reduce the error, this result highlights the potential for our approach to be trained with simulated data, then implemented to identify point locations in experimental data.

4. CONCLUSION

This work represents the first demonstration of machine learning applied to photoacoustic data visualization. We trained a CNN using simulated parameters (i.e. 2D point target positions, medium sound speeds, and target radii) to predict realistic image formation parameters directly from raw, pre-beamformed data. Our preliminary results demonstrate the feasibility of the proposed approach, and future work will test alternative CNNs that are trained to detect multiple point sources, partial wavefields, and identify artifacts for removal.

REFERENCES

[1] Muyinatu A Lediju Bell, Nathanael P Kuo, Danny Y Song, Jin U Kang, and Emad M Boctor. In vivo visualization of prostate brachytherapy seeds with photoacoustic imaging. *Journal of Biomedical Optics*, 19(12):126011–126011, 2014.

[2] Muyinatu A Lediju Bell, Xiaoyu Guo, Danny Y Song, and Emad M Boctor. Transurethral light delivery for prostate photoacoustic imaging. *Journal of Biomedical Optics*, 20(3):036002–036002, 2015.

[3] Muyinatu A Lediju Bell, Anastasia K Ostrowski, Ke Li, Peter Kazanzides, and Emad M Boctor. Localization of transcranial targets for photoacoustic-guided endonasal surgeries. *Photoacoustics*, 3(2):78–87, 2015.

[4] Blackberrie Eddins and Muyinatu A Lediju Bell. Design of a multifiber light delivery system for photoacoustic-guided surgery. *Journal of Biomedical Optics*, 22(4), 2017.

[5] Emma R Hill, Wenfeng Xia, Matthew J Clarkson, and Adrien E Desjardins. Identification and removal of laser-induced noise in photoacoustic imaging using singular value decomposition. *Biomedical Optics Express*, 8(1):68–77, 2017.

[6] Muyinatu A Lediju Bell, Nathanael Kuo, Danny Y Song, and Emad M Boctor. Short-lag spatial coherence beamforming of photoacoustic images for enhanced visualization of prostate brachytherapy seeds. *Biomedical Optics Express*, 4(10):1964–1977, 2013.

[7] Mithun Kuniyil Ajith Singh and Wiendelt Steenbergen. Photoacoustic-guided focused ultrasound (PAFUSion) for identifying reflection artifacts in photoacoustic imaging. *Photoacoustics*, 3(4):123–131, 2015.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[9] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *arXiv:1506.01497 [cs.CV]*, 2015.

[10] David Eigen, Christian Puhrsch, and Rob Fergus. Depth map prediction from a single image using a multi-scale deep network. In *arXiv:1406.2283 [cs.CV]*, 2014.

[11] M Nikoonahad and DC Liu. Medical ultrasound imaging using neural networks. *Electronics Letters*, 26(8):545–546, 1990.

[12] Bradley E Treeby and Benjamin T Cox. k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. *Journal of Biomedical Optics*, 15(2):021314–021314, 2010.

[13] Bradley E Treeby, Trond K Varslot, Edward Z Zhang, Jan G Laufer, and Paul C Beard. Automatic sound speed selection in photoacoustic image reconstruction using an autofocus approach. *Journal of Biomedical Optics*, 16(9):090501–090501, 2011.

[14] Neeraj Gandhi, Sungmin Kim, Peter Kazanzides, and Muyinatu A. Lediju Bell. Accuracy of a novel photoacoustic-based approach to surgical guidance performed with and without a da Vinci robot. *SPIE BiOS*, 2017.