A Machine Learning Method to Identify and Remove Reflection Artifacts in Photoacoustic Channel Data

Derek Allman*, Austin Reiter†, Muyinatu A. Lediju Bell*‡
*Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA
†Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA
‡Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

***Abstract*—Photoacoustic imaging is often used to visualize point-like targets, including circular cross sections of small cylindrical implants like brachytherapy seeds as well as circular cross sections of metal needles. When imaging these point-like targets in the presence of highly echogenic structures, the resulting image will suffer due to reflection artifacts which appear as true signals in the traditional beamformed image. We propose to use machine learning methods to identify these types of noise artifacts for removal. A deep convolutional neural network was trained to locate and classify source and reflection artifacts in photoacoustic channel data simulated in k-Wave. Simulated channel data contained one source and one artifact with varying target locations, medium sound speeds, and -3dB channel noise. In testing 3,998 simulated images, we achieved a 99.1% and 98.8% success rate in classifying sources and artifacts, respectively, while obtaining a misclassification rate below 3.1%, where a misclassification was defined as a source or artifact detected as an artifact or source, respectively. The network, which was only trained on simulated data, was then transferred to experimental data with 100% source classification accuracy and 0.40 mm mean source location accuracy. These results are promising as they show that a network trained with only simulated data can distinguish experimental sources and artifacts in photoacoustic channel data and display this information in a novel artifact-free image format.**

I. Introduction

Photoacoustic imaging is used to visualize optical properties through the preferential absorption of light [1], [2]. Structures in the body absorb the light which causes thermal expansion and creates a pressure wave that can be sensed with an ultrasound transducer and reconstructed (or beamformed) for image display. However, photoacoustic images are often degraded by strong acoustic reflections from hyperechoic structures. Traditional beamforming techniques involving time-of-flight models are not capable of reconstructing true acoustic sources when multiple scattering events are involved. Thus, reflections are mapped to incorrect locations in the resulting beamformed image. This can be problematic in clinical applications of photoacoustic imaging, where an image read by a clinician can potentially be incorrect or misleading.

There are several potential medical applications of photoacoustic imaging which are affected by these reflection artifacts. Such applications include brachytherapy for treatment of prostate cancer [3], in which brachytherapy seeds are visualized using photoacoustic imaging and the channel data is processed using either traditional delay-and-sum methods or short-lag spatial coherence imaging [4]. With both imaging methods it is difficult to differentiate between signals originating from the seed and those originating from reflection artifacts. Similarly, metal needles are ideal photoacoustic targets because of their high optical absorbance, but they are also affected by reflection artifacts (and increased background noise when applying filtering methods to reduce reflection artifacts) [5]. Reflections can also be caused tissue structures inside the subcutaneous fat layer [6].

PAFUSion [6], [7] uses ultrasound data to mimic the wavefields produced by photoacoustic sources and identify reflection artifacts for removal. However, this method makes the assumption of shared acoustic pathways for both ultrasound and photoacoustic data which is not always true. This method therefore has limited potential in a real-time environment (due to the requirement for matched ultrasound and photoacoustic images).

We propose to address outstanding challenges with reflection artifact reduction by employing deep convolutional neural networks (CNNs) [8]–[11]. CNNs have seen increasing popularity due to their success in modeling highly complex problems like those in image processing [8]. These networks can potentially be applied as an alternative to photoacoustic beamforming [12].

Bell and Reiter [12] showed that a deep network can be used to locate photoacoustic signals in channel data with an average positional accuracy of 0.28 mm and 0.37 mm in the depth and lateral image dimensions, respectively. Expanding on this work, we trained a deep neural network with simulated -3dB SNR photoacoustic channel data to locate and distinguish between sources and artifacts. We evaluated its performance on both simulated channel data with -3dB channel SNR and experimental channel data. Finally, we present a method for artifact removal utilizing the outputs of the CNN.

II. Methods

We trained a Faster-RCNN algorithm [11], VGG16 CNN network architecture [13] using photoacoustic channel data simulated with k-Wave [14]. Each image contained one 0.1 mm point source and one artifact. Photoacoustic sources were simulated with the range and increment size of our simulation

TABLE I: Range and Increment Size of Simulation Variables

Parameter	Min	Max	Increment
Depth Position (mm)	5	25	0.25
Lateral Position (mm)	5	30	0.25
Speed of Sound (m/s)	1440	1640	6

variables listed in Table I. As reflection artifacts tend to have a wave shape that is characteristic of signals at shallower depths, it is possible to simulate these artifacts by shifting a source signal deeper into the image. To generate the reflection artifacts, a photoacoustic source was shifted deeper in the image by the Euclidean distance, Δ, as described by the equation:

$$|\Delta| = \sqrt{(z_s - z_r)^2 + (x_s - x_r)^2} \tag{1}$$

where (x_s, z_s) are the 2D spatial coordinates of the source location and (x_r, z_r) are the 2D spatial coordinates of the physical reflector location.

To create a dataset large enough to train the network, point target locations were randomly selected from all possible source locations, while artifact locations were randomly selected from all possible points located less than 10 mm from the source. White Gaussian noise was added to each image at -3dB SNR, as most channel data contains some amount of noise. A total of 19,992 channel data images were synthesized for this simulated dataset, and 80% of the images were used for training while the remaining 20% of the images were used for testing.

The Faster R-CNN outputs consisted of the classifier prediction, corresponding confidence score (a number between 0 and 1), and the bounding box image coordinates for each detection. These detections were evaluated according to their classification results as well as their depth, lateral, and total positional errors.

Detections were classified as correct if the intersect-over-union (IoU) of the ground truth and detection bounding box was greater than 0.5 and their score was greater than an optimal value. This optimal value for each class and each network was found by first defining a line with a slope equal to the number of negative detections divided by the number of positive detections, where positive detections were defined as detections with a IoU greater than 0.5. This line was shifted from the ideal operating point (true positive rate of 1 and false positive rate of 0) down and to the right until it intersected with the receiver operating characteristics (ROC) curve. The point at which it first intersected with the ROC curve was determined to be the optimal score threshold. The ROC curve was created by varying the confidence threshold and plotting the rate of true and false positives at each tested threshold. The ROC curve indicates the quality of object detections made by the network. Misclassifications were defined to be a source detected as an artifact or an artifact detected as a source, and missed detections were defined as a source or artifact being detected as neither a source nor artifact. In addition to classification, misclassification, and missed detection rate, we also considered precision, recall, and area-under-curve (AUC) for the ROC curve.

We performed a controlled experiment to determine the feasibility of training with simulated data for the eventual identification and removal of artifacts in real data acquired from patients in a clinical setting. A 1 mm core diameter optical fiber was inserted in a needle and placed in the imaging plane between the transducer and a sheet of acrylic. This setup was placed in a water tank. The optical fiber was coupled to a Quantel (Bozeman, MT) Brilliant laser. When fired, the laser light from the fiber tip creates a photoacoustic signal in the water which propagates in all directions. This signal travels both directly to the transducer, creating the source signal, and to the acrylic which reflects the signal to the transducer, creating the reflection artifact. Seventeen channel data images were captured, each after changing the location of the transducer while keeping the laser and acrylic spacing fixed. The mean channel SNR of the experimental data was measured as 0.1dB and the artifacts were labeled by hand after observing the B-mode image.

After obtaining detection and classification results for the simulated and experimental data, we used the network outputs to display only the locations of the detected source signals in the image.

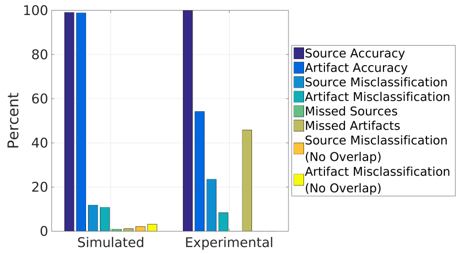
III. Results & Discussion

A. Classification Accuracy

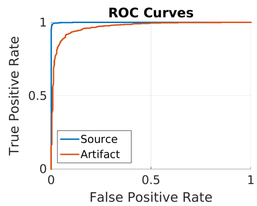
As shown in Fig. 1(a) our proposed network correctly classified sources and artifacts 99.1% and 98.9% of the time, respectively, and the misclassification rate was 11.8% and 10.8% for sources and artifacts, respectively, as shown in Fig. 1. Although these misclassification rates seem large, we noticed that there were several special cases where artifacts and sources overlapped, causing confusion in the source and artifact classifications. For example, artifacts that overlapped sources were detected and the corresponding bounding box was in the same location as the source bounding box, which caused an increase in our misclassification rates. We therefore took one additional step to report results after excluding the special overlapping cases from our analysis and obtained misclassification rates of 2.1% and 3.1% for sources and artifacts, respectively.

Fig. 1(b) shows the corresponding ROC curves for the sources and artifacts. The result for sources is shown in blue and the result for artifacts is shown in red with true positive rate on the vertical axis and false positive rate on the horizontal axis. In addition, precision, recall, and AUC all exceed 0.96. These factors confirm that our CNN is well suited to detect wavefronts in photoacoustic channel data as well as distinguish where they originated. While this work is primarily concerned with objects with circular cross-sections, this method can be extended to include objects with different shapes by training with different initial pressure distributions.

In addition to the simulation results, Fig. 1(a) shows that the percentage of correct source and artifact detections were 100% and 54.1%, respectively, for the 17 experimental images. The



(a)



(b)

Fig. 1: (a) Classification results. The bars from left to right show the accuracy of sources and artifact detections, the misclassification rate for sources and artifacts, the missed detection rate for sources and artifacts, and the misclassification rate for sources and artifacts after removing overlapping sources and artifacts from calculations. (b) Source and artifact ROC curves for simulated data.

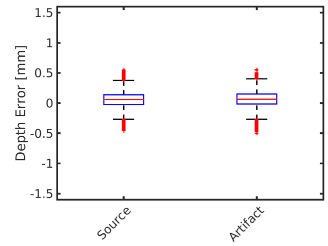
TABLE II: Summary of Classification Performance

Dataset	Source			Artifact		
	Precision	Recall	AUC	Precision	Recall	AUC
Simulated	0.9912	0.9910	0.9989	0.9744	0.9887	0.9689

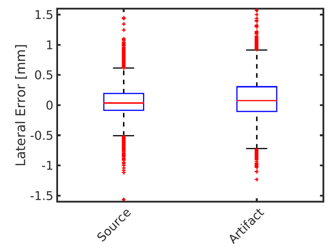
network provided more misclassifications (23.5% for sources and 8.3% for artifacts) when compared to its performance on simulated data (2.1% for sources and 3.1% for artifacts) due to lower confidence score. In addition, the artifact missed detection rate is higher at 45.83%, likely due to the presence of multiple reflections.. Despite these large errors, we will show in Section III.C that the source accuracy result is of primary interest for our chosen artifact removal method. Therefore, this result indicates that a network trained with only simulated channel data can be transferred to experimental data while maintaining a high level of performance for source detection and classification.

B. Location Errors

The box-and-whiskers plots in Fig. 2 demonstrate the depth and lateral errors for sources and artifacts for the simulated data. 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. Fig. 2(a) show that the networks are more accurate in the depth dimension, where errors (including outliers) were frequently less than 0.6 mm, when compared to errors in the lateral dimension (Fig. 2(b)), where outliers were as large as 1.5-2.0 mm. However, in both cases, the median values were consistently less than 0.1-0.5 mm. These results indicate that



(a)



(b)

Fig. 2: Summary of distance errors for all tested simulated data in the depth (a) and lateral (b) dimensions for sources and artifacts. Note that our depth errors are consistently lower than our lateral errors.

in addition to providing excellent source classification results, our network also provides accurate position estimates of true source locations.

C. Artifact Removal

Fig. 3 shows the results of our method for removing artifacts. Sample channel input data to the network is shown in Fig. 3 (a), and the corresponding B-mode image is shown in Fig. 3 (b). To display objects which were classified as sources (Fig. 3 (e)), a disc-shaped object was placed at the center of the detected bounding box and displayed with a diameter of $\pm 2\sigma$, where σ refers to the standard deviation of the location errors shown in Fig. 2.

Each experimental image had one source signal and at least one reflection artifact, as shown in Fig. 3(d). We know that only one of these signals is a true source, because we only had one source in the image. The corresponding beamformed image Fig. 3(e) shows the problematic reflection artifact being removed in Fig. 3(f). We additionally note that several of the experimental images contained multiple reflection artifacts due to reverberations from the walls of the water tank, as seen in Fig. 3(g), yet these multiple artifacts, clearly present in the beamformed image (Fig. 3(h)), do not affect our artifact removal method (Fig. 3(i)).

In these three cases (Figs. 3(c), 3(f) and 3(i)), one major benefit of our display method is that we can visualize true sources with an arbitrarily high contrast. The new image is not corrupted by reflection artifacts because we do not display them, and we therefore achieve noise-free, high resolution images of the original point target.

IV. Conclusion

We trained a CNN using simulated images of photoacoustic channel data and showed that the network can distinguish between a simulated source and artifact in the presence of

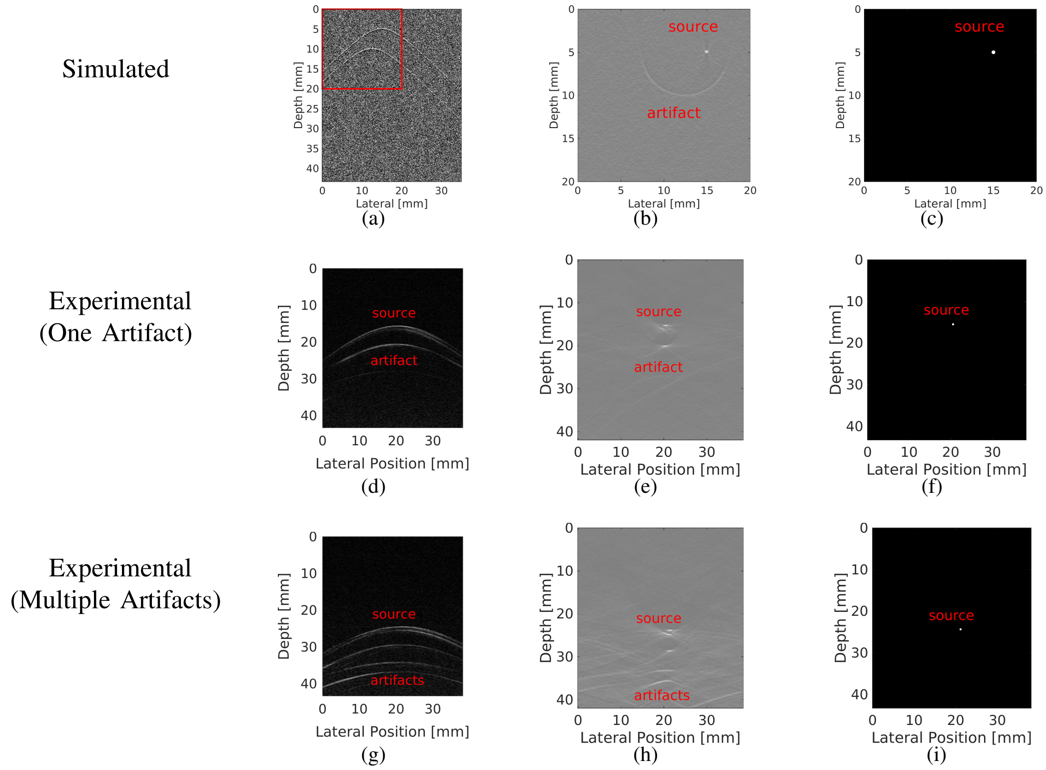


Fig. 3: Sample images from channel data (a) before and (b) after applying traditional beamforming. Artifact removal with a (c) CNN-based image that displays the location of the detected source based on the location of the bounding box. The red box in (a) indicates the portion of the images displayed in (b,c). (d) Sample image of experimental channel data containing one source and an associated reflection artifact. (e) Corresponding beamformed image. (f) Corresponding image created with our CNN-based artifact removal method. (g) Sample image of experimental channel data containing one source and multiple reflection artifacts. (h) Corresponding beamformed image (i) Corresponding image created with our CNN-based artifact removal method.

channel noise. We also demonstrated that we can accurately locate sources and artifacts in simulated images. In addition, our network was successfully transferred to experimental data without any additional training and achieved similar classification performance when detecting true sources. This approach highlights the potential for elimination of reflection artifacts for interventional photoacoustic images and a similar concept could potentially be applied to improve ultrasound image quality.

References

[1] P. Beard, "Biomedical photoacoustic imaging," *Interface focus*, p. rsfs20110028, 2011.

[2] M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Review of scientific instruments*, vol. 77, no. 4, p. 041101, 2006.

[3] M. A. Lediju Bell, N. P. Kuo, D. Y. Song, J. U. Kang, and E. M. Boctor, "In vivo visualization of prostate brachytherapy seeds with photoacoustic imaging," *Journal of Biomedical Optics*, vol. 19, no. 12, p. 126011, 2014. [Online]. Available: http://dx.doi.org/10.1117/1.JBO.19.12.126011

[4] M. A. L. Bell, N. Kuo, D. Y. Song, and E. M. Boctor, "Short-lag spatial coherence beamforming of photoacoustic images for enhanced visualization of prostate brachytherapy seeds," *Biomedical optics express*, vol. 4, no. 10, pp. 1964–1977, 2013.

[5] J. Su, A. Karpiouk, B. Wang, and S. Emelianov, "Photoacoustic imaging of clinical metal needles in tissue," *Journal of biomedical optics*, vol. 15, no. 2, pp. 021 309–021 309, 2010.

[6] M. K. A. Singh, M. Jaeger, M. Frenz, and W. Steenbergen, "In vivo demonstration of reflection artifact reduction in photoacoustic imaging using synthetic aperture photoacoustic-guided focused ultrasound (pafusion)," *Biomedical optics express*, vol. 7, no. 8, pp. 2955–2972, 2016.

[7] M. K. A. Singh and W. Steenbergen, "Photoacoustic-guided focused ultrasound (pafusion) for identifying reflection artifacts in photoacoustic imaging," *Photoacoustics*, vol. 3, no. 4, pp. 123–131, 2015.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.

[10] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.

[12] A. Reiter and M. A. L. Bell, "A machine learning approach to identifying point source locations in photoacoustic data," in *Proc. of SPIE Vol*, vol. 10064, 2017, pp. 100 643J–1.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR), 2015*, 2014.

[14] B. E. Treeby and B. T. Cox, "k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave-fields," *J. Biomed. Opt.*, vol. 15, no. 2, p. 021314, 2010.