Learning from Skinless Datasets to Improve Target Detection in Photoacoustic Breast Imaging Across Skin Tones

Julia Urszula Drabek[1*], Madhavi Tripathi[1], Jamie Enslein[1], Kelly Myers[2], Babita Panigrahi[2], Muyinatu A. Lediju Bell[1,3,4,5*]
[1]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland, USA
[2]Department of Radiology, Johns Hopkins Medicine, Baltimore, Maryland, USA
[3]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, USA
[4]Department of Computer Science, Johns Hopkins University, Baltimore, Maryland, USA
[5]Department of Oncology, Johns Hopkins Medicine, Baltimore, Maryland, USA

ABSTRACT

Photoacoustic imaging is promising to detect breast cancer. However, darker skin tones are known to absorb more light, increasing acoustic clutter. We developed a conditional generative adversarial network (cGAN) that learns from simulated channel data from targets located within skinless breast models to improve target detectability in channel data acquired from darker skin tones. Training of the cGAN was performed using k-Wave-simulated photoacoustic channel data. Ground truth channel data were generated by replacing the skin layer with water to eliminate the optical absorption effects of skin. The trained cGAN was then tested on *in vivo* data from a 75-year-old woman (breast tissue individual typology angle (ITA)=–9°, areolar tissue ITA=-72°), scheduled for ultrasound-guided core-needle breast biopsy. When applied to simulated data, the mean generalized contrast-to-noise ratio of channel data ($gCNR_{ch}$) from the output improved by 3-64% for skin tones ranging dark to very light, relative to the mean $gCNR_{ch}$ of input channel data. The mean generalized contrast-to-noise ratio (gCNR) of SLSC images created from channel data output by the cGAN increased relative to the mean gCNR of SLSC images created from channel data input to the cGAN by 6-43% for skin tones ranging dark to tan. When applied to the *in vivo* case, $gCNR_{ch}$ increased from 0.13 to 0.39 and gCNR increased from 0.95 to 0.96 This initial demonstration of feasibility is generally promising to mitigate melanin-induced artifacts and enhance target detectability in photoacoustic breast images.

1. INTRODUCTION

Photoacoustic imaging is a promising modality for breast cancer detection,[1] integrating the spatial resolution of ultrasound with optical imaging contrast.[2–4] However, skin tone impacts photoacoustic image quality,[5,6] as increased melanin content reduces optical penetration and degrades image quality.[5,6] Fernandes *et al.*[7] identified the skin as a primary source of acoustic clutter in fast Fourier transform (FFT) reconstructions of simulated breast tissue. These simulations were furher developed by Rasquinha *et al.*[5,8] to quantify the image quality degradation associated with increasing melanin content for conventional amplitude-based reconstruction methods (e.g., FFT-based reconstruction and delay-and-sum (DAS) beamforming).

Short-lag spatial coherence (SLSC) beamforming was previously demonstrated to reduce skin tone bias in photoacoustic imaging.[5,9] SLSC beamforming is advantageous in this regard because the technique computes the spatial coherence of signals received by transducer elements, rather than relying solely on signal amplitude as in DAS or FFT-based methods,[10,11] which effectively enables clutter suppression and contrast enhancement.[12–14]

In addition to beamforming methods, machine learning has been leveraged to improve image quality and enable target visualization.[15–17] Else *et al.*[15] developed a gradient-boosted regression tree model for spectral unmixing to reduce photoacoustic oximetry errors in individuals with darker skin. Huang *et al.*[16] developed a deep learning approach to segment the skin layer, then applied adaptive weighting to compensate for signal

*Address correspondence to: jdrabek1@jh.edu and mledijubell@jhu.edu

Photons Plus Ultrasound: Imaging and Sensing 2026, edited by Alexander A. Oraevsky, Lihong V. Wang,
Proc. of SPIE Vol. 13851, 138510N · 2026 Published by SPIE · 0277-786X · doi: 10.1117/12.3083562

loss beneath the skin and enhance deeper structures. While these approaches are promising, they operate on beamformed data, which potentially limits the maximum possible improvements.

Rather than operating on beamformed data, we believe that operations applied to channel data has the potential to maximize pattern inference and further enhance machine learning capabilities. In particular, training a deep neural network (DNN) to learn image features from raw channel data, with no skin-related artifacts present, is expected to enable the DNN to infer an ideal artifact-free appearance. The resulting output channel data can then be converted to an image with any beamformer.

To investigate the potential of our approach, we first train with simulated data containing no skin (i.e., skin-removed data), then apply our trained DNN to skin-present simulation and *in vivo* data. Considering the previously reported success with the SLSC beamformer,[5,7] we use this approach to evaluate photoacoustic image quality before and after intervention from our trained DNN.

2. METHODS

2.1 DNN Architecture & Training

Our DNN is a conditional generative adversarial network (cGAN) consisting of a a U-Net-inspired[18] generator and a PatchGAN-inspired[19] discriminator. The generator used an encoder-decoder structure with skip connections to combine semantic and spatial information. The discriminator evaluated the realism of the data, producing patch-wise probabilities indicating the likelihood of each region being real. This cGAN was trained to accept channel data with a skin layer present (i.e., skin-present data) and output ground truth channel data without the skin layer (i.e., skin-removed data). Training was performed end-to-end in Pytorch, using the Adam optimization algorithm for 50 epochs, with a batch size of 4 and a learning rate parameter of 10^{-4}. The training hardware consisted of four Nvidia Quadro RTX 6000 GPUs, each with 24 GB VRAM and an Intel Core-i7 7820X processor.

2.2 Simulated Data for Training and Testing

Skin-present simulation data were generated following the method described by Rasquinha *et al.*[5] In particular, a simulation volume, with breast tissue classified as extremely dense, was constructed based on a realistic 3D breast model immersed in water.[20] Spherical inclusions were placed within the volume at the axial depths and lateral positions indicated in Table 1. MCXLab[21] was employed to compute optical fluence within the simulated volumes, with optical properties corresponding to an 800 nm laser wavelength. To simulate the various skin tone categories defined in,[22–24] the ITA values of the skin layer was varied, as reported in Table 1.

To create skin-removed simulation volumes, the process described by Fernandes *et al.*[7] was employed. In particular, the skin layer in each volume created in the skin-present dataset was replaced with water. As a result, the effects of ITA variations and optical properties of skin were not included in the associated MCXLab optical

Table 1. Simulation variables, total number of simulated volumes, and total number of channel data files in skin-present and skin-removed datasets.

Parameter	**Value**	**Increment**	**Total**
Target diameter [mm]	0.5 to 3	0.25	11
Target depth [mm]	5 to 13	1	9
Target lateral position [mm]	45	N/A	1
ITA [°]	-54 to 60	2	58
Skin-present simulations			
Total number of volumes			99
Total number of channel data files			5,742
Skin-removed simulations			
Total number of volumes			99
Total number of channel data files			99

simulations. The resulting optical fluence was utilized to generate raw channel data using the k-Wave toolbox.[25] Noise corresponding to a channel signal-to-noise ratio (SNR) of 20 dB was added to the channel data. The total number of channel data files are reported in Table 1. Each total of skin-removed and skin-present channel data files were split 80% for training and 20% for testing.

During training, each simulated skin-present file was paired with a corresponding simulated skin-removed file, matched by target size and depth. To maintain a lightweight, proof-of-concept approach, training was limited to a single wavelength. Prior to cGAN input, channel data were resized to 512×512 and normalized to the brightest pixel in each image.

2.3 *In Vivo* Test Data

As an additional test case, *in vivo* photoacoustic radiofrequency channel data were acquired from a 75-year-old woman scheduled for ultrasound-guided core-needle breast biopsy, using an Imagio® ultrasound/photoacoustic imaging system (Seno Medical, Inc., San Antonio, TX, USA). There were 256 receive scan lines per image. Signals received by adjacent elements were summed prior to data storage, resulting in 128 channels for beamforming. The transmitted optical wavelength of this test data was 1064 nm. ITA was measured with a colorimeter (Delfin Technologies, Kuopio, Finland), resulting in -9° on breast tissue and -72° on the areola. This prospective data acquisition for the Advanced Ultrasound Signal Processing of Suspicious Breast Images (AUSPICIOUS) observational study (NCT07206888) occured after informed consent and approval from the Johns Hopkins Institutional Review Board (Protocol No. IRB00127110).

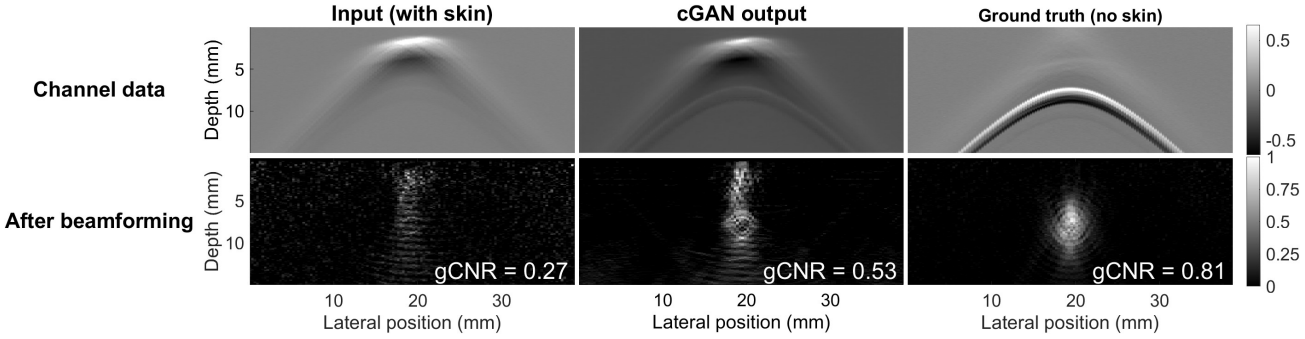
2.4 Evaluation Metrics

To assess target detectability before and after applying the cGAN, the generalized contrast-to-noise ratio (gCNR) was computed before and after applying short-lag spatial coherence (SLSC) beamforming. The calculations performed before beamforming were applied to the raw channel data, resulting in gCNR_{ch} values.[26] Given differences in *a priori* information about target locations, two independent region of interest (ROI) creation processes were employed to calculate the gCNR_{ch} and gCNR of simulated and *in vivo* data. For simulated data, a circular target ROI was placed at the simulated target depth in beamformed images, with a radius matching the target size, and a background ROI was defined as an annulus centered at the same location. The inner radius of the annulus was $\sqrt{2}$ times larger than the target radius, with equal area to that of the target ROI. The channel data target ROI was bounded by two hyperbolic curves,[27] with the same-shaped background ROI shifted by 10 mm. For beamformed *in vivo* data, ellipses were fit to segmented regions after thresholding amplitudes ≥ 0.8 to create target ROIs, which were laterally shifted by 10 mm to create corresponding background ROIs. The target ROI locations and sizes were then employed to create hyperbolic curves[27] with the same-shape background ROI shifted axially 10 mm to calculate the gCNR_{ch} of *in vivo* data.

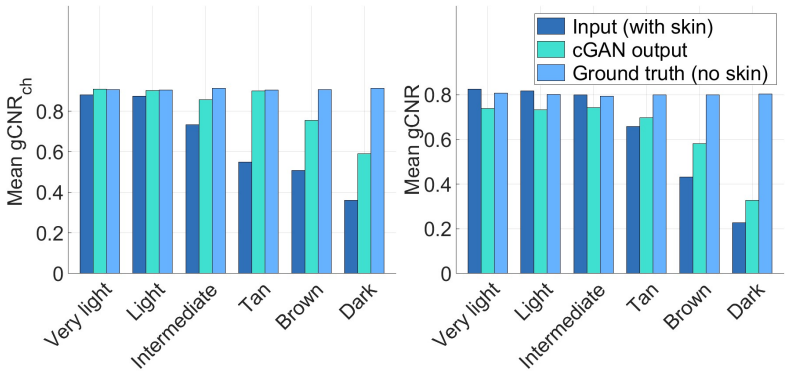
3. RESULTS

Fig. 1 shows simulation results obtained with the cGAN. In Fig. 1(a), a simulated 1.75 mm-diameter target underlies brown skin (ITA = –28°) at a depth of 8 mm. After cGAN processing, the gCNR_{ch} increased form 0.04 to 0.92, and the gCNR increased from 0.27 to 0.53. As expected, the targets are best visualized in the ground truth, skin-removed data. In Fig. 1(b), the mean gCNR_{ch} and gCNR of input, cGAN output, and ground truth data are plotted as a function of the six ITA-based skin tone categories simulated. There is a progressive decline in the input and output gCNR_{ch} and gCNR as skin tone darkens. Although the cGAN does not fully recover the target values achieved with no skin present, the mean gCNR_{ch} and gCNR of the cGAN output is consistently greater than that of the input with skin, resulting in 3-64% gCNR_{ch} improvements and 6-43% gCNR improvements across skin tones.

Fig. 2 shows an example result obtained with the patient data. After cGAN processing, the gCNR_{ch} increased from 0.13 to 0.39, and the gCNR of the beamformed target in the image increased from 0.95 to 0.96.



(a)



(b)

Figure 1. (a) Example simulation result with brown skin (ITA -28°). The mean of channel data was set to zero before display. (b) Comparison of mean gCNR$_{ch}$ and gCNR across different skin tone categories.

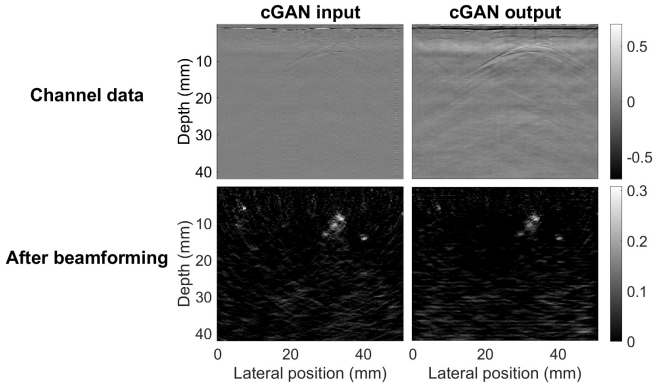


Figure 2. Example *in vivo* result with ITA -72°. The mean of channel data was set to zero before display.

4. DISCUSSION

This paper introduces a cGAN that mitigates disparities related to skin tone in photoacoustic breast images. Training the cGAN on skin-removed simulated channel data enabled the recovery of low-amplitude signals to improve both target detectability in channel data and beamformed image quality (Fig. 1). It is additionally promising that the cGAN effectively generalized beyond the controlled simulation environment, maintaining an improvement in image quality for the *in vivo* example (Fig. 2).

The cGAN appears to work by enhancing the detectability of waveforms in the channel data. In particular, the increase in gCNR$_{ch}$ after cGAN processing indicates an enhancement of the target waveform relative to surrounding noise and clutter. After beamforming, the cGAN improvements to the channel data successfully enhance simulated target signals and reduce the background clutter, as indicated by the gCNR values.

One limitation of this work is that the cGAN was tested on a single *in vivo* example. Future work will evaluate the cGAN on a larger and more diverse set of clinical data with a wider range of skin tones. The impact on various image reconstruction methods may also be investigated.

5. CONCLUSION

This paper introduces a novel application of cGANs to address skin tone challenges in photoacoustic breast imaging. We demonstrated that cGAN preprocessing improves gCNR_{ch} and gCNR in simulated data. Therefore, this type of channel data correction has the potential to improve downstream performance, underscoring future promise for clinical translation.

ACKNOWLEDGMENTS

This work is supported by NIH R01 EB032960 and grant number 2022-309513 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation. We are additionally grateful for the in-kind support of Seno Medical, Inc.

REFERENCES

1. N. Nyayapathi and J. Xia, "Photoacoustic imaging of breast cancer: A mini review of system design and image features," *Journal of Biomedical Optics* **24**(12), pp. 121911–121911, 2019.
2. M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Review of Scientific Instruments* **77**, p. 041101, Apr. 2006.
3. P. Beard, "Biomedical photoacoustic imaging," *Interface focus* **1**(4), pp. 602–631, 2011.
4. J. Laufer, "Photoacoustic imaging: Principles and applications," in *Quantification of Biophysical Parameters in Medical Imaging*, I. Sack and T. Schaeffter, eds., pp. 345–363, Springer International Publishing, 2024.
5. R. D. Rasquinha, M. R. Gubbi, and M. A. L. Bell, "Impact of skin tone on target size detectability in photoacoustic breast imaging," *Biophotonics Discovery* **2**(1), pp. 012502–012502, 2025.
6. Y. Mantri and J. V. Jokerst, "Impact of skin tone on photoacoustic oximetry and tools to minimize bias," *Biomedical Optics Express* **13**, pp. 875–887, Feb 2022.
7. G. S. Fernandes, T. Z. Pavan, and M. A. L. Bell, "In silico demonstrations of the impact of wavelength and skin tone on photoacoustic breast imaging," in *Proceedings of the IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, IEEE, 2023.
8. R. D. Rasquinha, M. R. Gubbi, and M. A. L. Bell, "Influence of skin tone on target size detectability in photoacoustic breast imaging," in *Proceedings of SPIE Photonics West*, SPIE, 2025.
9. G. S. Fernandes, J. H. Uliana, L. Bachmann, A. A. Carneiro, M. A. L. Bell, and T. Z. Pavan, "Mitigating skin tone bias in linear array in vivo photoacoustic imaging with short-lag spatial coherence beamforming," *Photoacoustics* **33**, p. 100555, 2023.
10. M. A. Lediju, G. E. Trahey, B. C. Byram, and J. J. Dahl, "Short-lag spatial coherence of backscattered echoes: Imaging characteristics," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **58**(7), pp. 1377–1388, 2011.
11. M. T. Graham and M. A. L. Bell, "Photoacoustic spatial coherence theory and applications to coherence-based image contrast and resolution," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **67**(10), pp. 2069–2084, 2020.
12. 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* **4**(10), pp. 1964–1977, 2013.
13. M. A. L. 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* **19**(12), p. 126011, 2014.
14. M. A. L. Bell, X. Guo, H. J. Kang, and E. Boctor, "Improved contrast in laser-diode-based photoacoustic images with short-lag spatial coherence beamforming," in *Proceedings of the IEEE International Ultrasonics Symposium (IUS)*, pp. 37–40, IEEE, 2014.

15. T. R. Else, L. Hacker, J. Gröhl, E. V. Bunce, R. Tao, and S. E. Bohndiek, "Effects of skin tone on photoacoustic imaging and oximetry," *Journal of Biomedical Optics* **29**(S1), pp. S11506–S11506, 2024.
16. C. Huang, E. Zheng, W. Zheng, H. Zhang, Y. Cheng, X. Zhang, V. Shijo, R. W. Bing, I. Komornicki, L. M. Harris, *et al.*, "Enhanced clinical photoacoustic vascular imaging through a skin localization network and adaptive weighting," *Photoacoustics* **42**, p. 100690, 2025.
17. N. Wang, T. Chen, C. Liu, and J. Meng, "Intelligent skin-removal photoacoustic computed tomography for human based on deep learning," *Journal of Biophotonics* **17**(10), p. e202400197, 2024.
18. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 18th International Conference*, **III**, pp. 234–241, Springer, (Munich, Germany), Oct. 5–9 2015.
19. C. Li and M. Wand, "Precomputed real-time texture synthesis with markovian generative adversarial networks," in *Computer Vision – ECCV 2016, 14th European Conference*, **III**, pp. 702–716, Springer, (Amsterdam, The Netherlands), Oct. 11-14 2016.
20. Y. Lou, W. Zhou, T. P. Matthews, C. M. Appleton, and M. A. Anastasio, "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *Journal of biomedical optics* **22**(4), pp. 041015–041015, 2017.
21. Q. Fang and D. A. Boas, "Monte carlo simulation of photon migration in 3d turbid media accelerated by graphics processing units," *Optics express* **17**(22), pp. 20178–20190, 2009.
22. B. C. K. Ly, E. B. Dyer, J. L. Feig, A. L. Chien, and S. Del Bino, "Research techniques made simple: cutaneous colorimetry: a reliable technique for objective skin color measurement," *Journal of Investigative Dermatology* **140**(1), pp. 3–12, 2020.
23. A. Chardon, I. Cretois, and C. Hourseau, "Skin colour typology and suntanning pathways," *International journal of cosmetic science* **13**(4), pp. 191–208, 1991.
24. S. Del Bino, J. Sok, E. Bessac, and F. Bernerd, "Relationship between skin response to ultraviolet exposure and skin color type," *Pigment cell research* **19**(6), pp. 606–614, 2006.
25. B. E. Treeby and B. T. Cox, "k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics* **15**(2), pp. 021314–021314, 2010.
26. M. R. Gubbi and M. A. L. Bell, "Deep learning-based photoacoustic visual servoing: Using outputs from raw sensor data as inputs to a robot controller," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 14261–14267, IEEE, 2021.
27. M. R. Gubbi and M. A. L. Bell, "Deep learning to localize photoacoustic sources in three dimensions: Theory and implementation," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **72**(6), pp. 786–805, 2025.