Application of CohereNet to Photoacoustic Data for Non-Invasive, *In Vivo*, Subcutaneous Imaging

José Timaná[1], Guilherme S. P. Fernandes[2], Theo Z. Pavan[2], Muyinatu A. Lediju Bell[3,4,5]
[1]Department of Engineering, Pontificia Universidad Católica del Perú, Perú
[2]Department of Physics, FFCLRP, University of São Paulo, Brazil
[3]Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD
[4]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD
[5]Department of Computer Science, Johns Hopkins University, Baltimore, MD

2023 IEEE International Ultrasonics Symposium (IUS) | 979-8-3503-4645-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/IUS51837.2023.10306424

Abstract*—Short-lag spatial coherence (SLSC) beamforming of photoacoustic signals reduces acoustic clutter and enhances the contrast of underlying signals of interest. However, the original SLSC imaging algorithm is also known to be computationally expensive to implement. CohereNet, a custom deep neural network (DNN), was previously introduced to address the computational complexity of SLSC beamforming when applied to ultrasound data. Although it is conceivable that CohereNet can be translated to photoacoustic data, this translation is challenged by differences in data patterns and imaging depths when compared to ultrasound data, and it requires a similar diversity of photoacoustic training data to that present in ultrasound training data. Therefore, we propose to translate CohereNet to photoacoustic imaging after taking advantage of the diversity of signal-to-noise ratios (SNRs) available with transcutaneous *in vivo* data containing various levels of acoustic clutter. Raw channel data were acquired from the distal forearm of 18 volunteers with a variety of skin constitutive pigmentation, with 12 volunteers designated for training, 3 for validation, and the remaining 3 for testing. Time-delayed kernels and corresponding spatial coherence functions were computed and utilized as DNN training inputs and outputs, respectively. The correlation between standard central processing unit (CPU) SLSC and DNN SLSC images in the test set were 0.76, 0.78, and 0.84 when acquired with 750 nm, 810 nm, and 870 nm wavelengths, respectively. In addition, a mean SNR improvement of 2.5 dB was achieved with DNN SLSC imaging (relative to CPU SLSC imaging), due to the smoother coherence functions created with CohereNet. Results highlight the versatility of CohereNet, particularly when translating from ultrasound to photoacoustic data after appropriate steps are taken to achieve success, with no evidence of overfitting to training data.

I. Introduction

Photoacoustic imaging is an emerging non-invasive modality that offers precise visualization of vascular structures, oxygen saturation levels, and pathologies due to its sensitivity to the optical properties of tissues [1], [2]. Among photoacoustic beamforming techniques, short-lag spatial coherence (SLSC) beamforming, has pronounced proficiency to enhance image contrast and reduce acoustic clutter [3]–[6]. Most recently, SLSC has also been demonstrated to reduce skin tone bias in photoacoustic images [7].

CohereNet [8], a custom fully connected deep neural network (DNN), was recently introduced as a method to directly compute a coherence function from time-delayed channel data. This approach enables faster and more efficient image reconstruction with greater fidelity to the standard central processing unit (CPU) algorithm when compared to the shortcuts required to implement the graphical processing unit (GPU)-based SLSC approach [9]. While previous training approaches were focused on ultrasound data [8], [10], the prospect of applying CohereNet to photoacoustic SLSC beamforming using transfer learning is promising.

However, the translation of CohereNet from ultrasound to photoacoustic SLSC beamforming encounters three specific challenges. First, different data patterns arise due to the inherent characteristics of the different imaging processes (e.g., reflectivity of abundant subresolution and subcellular scattering structures from tissue in ultrasound imaging vs. optical absorption and conversion to acoustic energy of more isolated photoacoustic targets). Second, non-invasive photoacoustic images are constrained by limited optical penetration depths [11], thereby operating at more superficial depths when compared to the deeper probing capabilities of ultrasound. Third, to achieve success with ultrasound data, CohereNet was trained with highly heterogenous *in vivo* breast data [8], which offers multiple examples of expectations for coherence functions in a single transmit-receive data acquisition sequence.

We propose to translate CohereNet to photoacoustic data by training with photoacoustic data that contains a diversity of signal-to-noise ratios (SNRs) and clutter levels when imaging through the skin. This new training dataset is anticipated to overcome challenges with differences in data patterns and imaging depths when compared to a network exclusively trained with ultrasound data. In addition, the desired photoacoustic data diversity is possible by utilizing a previously acquired dataset containing various skin tones [7].

II. Methods

A photoacoustic imaging system comprising an Nd:YAG laser (Brilliant B, Quantel Laser) connected to an OPO (MagicPRISM, Opotek) and a commercial ultrasound machine (SonixOP, Ultrasonix) with a parallel acquisition module (SonixDAQ, Ultrasonix) was utilized in this study. The laser was coupled with a trifurcated optical fiber bundle (77536, Newport), and the ultrasound system was attached to a linear array transducer with 128 receiving elements (L14-5/38, Ultrasonix) to acquire photoacoustic channel data from the distal forearm of 18 volunteers with different skin pigmentation.

For each volunteer, 2-4 frames of photoacoustic channel data were acquired with optical wavelengths of 750 nm, 810 nm, and 870 nm. These data were initially acquired and utilized for a previous study [7], wherein skin tones were quantitatively classified using the Individual Typology Angle (ITA°) measured using a colorimeter (Delta Vista 450G). Lower ITA° values indicate higher melanin concentration, which results in increased acoustic clutter [7], [12].

SLSC beamforming [13] calculates the normalized cross-correlation function between the time-delayed signals received by different elements of an array as a function of spatial lag, defined and implemented as follows on a CPU:

$$\hat{R}(m) = \frac{1}{N-m} \sum_{i=1}^{N-m} \frac{\sum_k s_{k,i}(n) s_{k,i+m}(n)}{\sqrt{\sum_k s_{k,i}^2(n) \sum_k s_{k,i+m}^2(n)}} \quad (1)$$

where m is the lag in the number of elements, N is the number of transducer receiving elements, and $s_{k,i}(n)$ is a time-delayed, zero-mean kernel consisting of k axial samples, each received at element i and centered at depth n. The resulting spatial coherence function is summed up to a specific short-lag value, M, as follows:

$$R_{sl} = \int_1^M \hat{R}(m) dm \approx \sum_{m=1}^{M} \hat{R}(m), \quad (2)$$

yielding the value of the SLSC pixel. This process is repeated for each lateral and axial position in the image. Therefore, one axial kernel of photoacoustic channel data $s_{k,i}(n)$ consists of k axial samples, each measured across the entire receive aperture and centered at depth n.

To implement the DNN approach to SLSC image formation, photoacoustic channel data were formatted to consist of 7 axial samples to form the input $s_{k,i}(n)$, which yielded an output of $\hat{R}(m)$ for each element i and depth n. Although CohereNet was designed to accept 64 elements [8], the photoacoustic channel data in this study were acquired with 128 elements. Therefore, the input $s_{k,i}(n)$ was resized using the *imresize* MATLAB function (MathWorks Inc, Natick, MA) with a bicubic kernel to obtain 7×64 inputs from the original 7×128 matrix. Opting for the *imresize* function over straightforward downsampling was driven by the necessity to respect the continuous and intricately connected characteristics of the photoacoustic pressure wave. Bicubic interpolation considers values from several neighboring channels, safeguarding the original coherence of the waveform and the subtle characteristics embedded within. In addition, given that the network architecture inherently expects an output shape of 1x64, we chose to use only the initial 64 out of the possible 127 lags, $\hat{R}(m)$, from our 128-element data. This approach aligns the photoacoustic data with the architecture and structure of CohereNet, thus streamlining the transfer learning process without major model modifications.

CohereNet was trained using empirically optimized hyper-parameters, including a batch size of 256, a Nadam optimizer, 20 epochs, and a learning rate of 0.00005. This training was implemented using Keras [14] with a Tensorflow backend [15].

TABLE I
NUMBER OF UNIQUE VOLUNTEERS IN THE TRAINING, VALIDATION, AND TESTING DATA SETS PER SKIN TONE CATEGORY

Skin tone category	Number of unique volunteers		
	Training set	Validation set	Test set
Light	1	1	-
Intermediate	3	-	1
Tan	3	1	1
Brown	3	1	-
Dark	2	-	1

A total of 3.3, 0.6, and 0.8 million samples were included in the training, validation, and test sets, respectively. Each set included at least three different skin tone classifications, with additional details regarding the number of unique volunteers for each skin tone reported in Table I.

Quantitative comparisons between CPU and DNN SLSC images were performed using image-to-image correlation coefficients and the SNR of identified blood vessels. SNR is defined as follows:

$$\mathrm{SNR} = \frac{S_i}{\sigma_o}, \quad (3)$$

where S_i and σ_0 represent the mean and standard deviation, respectively, of a ROI within and outside of the vessel, respectively. The ROIs used to obtain S_i and σ_0 were located at the same depth for each vessel. These values were calculated before applying log-compression. In addition, a novel overfitting detection method introduced by Zhang *et al.* [16] was employed. This technique uses three types of artificial channel data (two binary images and an image of Gaussian random noise) to quickly discern overfitting in DNNs.

III. RESULTS AND DISCUSSION

Fig. 1 shows representative examples of normalized spatial coherence functions obtained from a vessel location. The DNN output achieves a smoother profile when compared to the CPU result. This smoothing ability was similarly observed when CohereNet was trained and tested with ultrasound data [8].

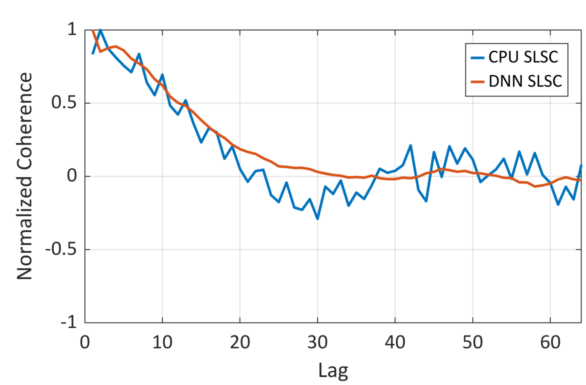

Fig. 1. Normalized coherence function example at one SLSC pixel location within a blood vessel.

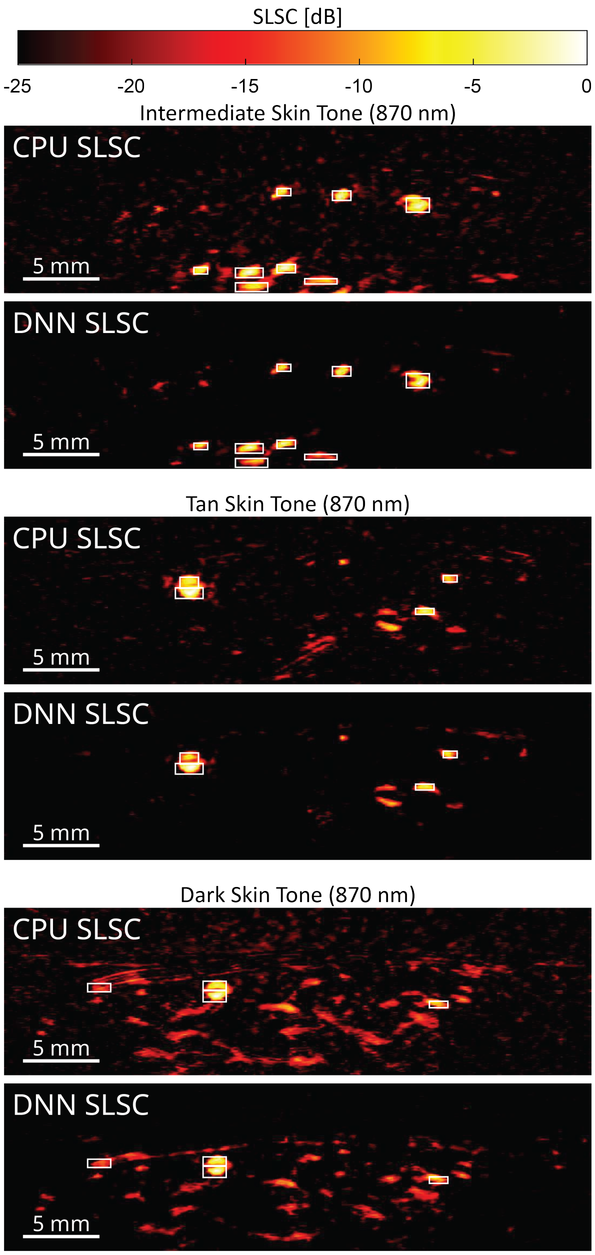

Fig. 2. Test set CPU and DNN SLSC images acquired with an optical wavelength of 870 nm. Vessels are highlighted by white the boxes, which denote regions of interest when calculating SNR.

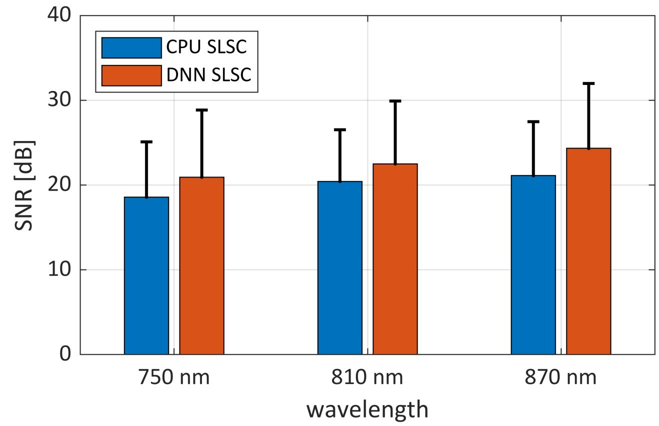

Fig. 3. Mean $\pm$ one standard deviation of SNR measurements as a function of wavelength, calculated for the three volunteers in the test set.

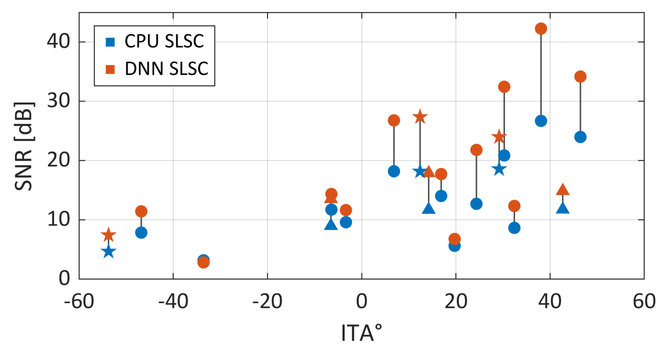

Fig. 4. Mean SNR calculated for the 18 volunteers compared to ITA° values (●: training set, ▲: validation set, ★: test set).

Fig. 2 shows example SLSC images created without and with CohereNet (i.e., CPU SLSC and DNN SLSC images, respectively) when inputting the raw channel data from three volunteers with skin tones previously classified as intermediate, tan, and dark [7], [17], [18]. Vessels highlighted by white boxes denote regions of interest for SNR calculations. The smoother coherence function outputs achieved with CohereNet leads to a qualitative reduction of acoustic clutter in the image background, when compared to corresponding CPU SLSC images. The test set produced mean image-to-image correlation values of 0.76, 0.78, and 0.84 when data were acquired with 750 nm, 810 nm, and 870 nm wavelengths, respectively.

Fig. 3 shows the mean $\pm$ one standard deviation of SNR measurements as a function of wavelength for the CPU SLSC and DNN SLSC results. These values were computed based on the data obtained from the volunteers included in the test set. An average SNR improvement of 2.5 dB across the three wavelengths was achieved with DNN SLSC imaging relative to CPU SLSC imaging. This enhancement is attributed to the reduced standard deviation in the background, stemming from the smoother DNN coherence function outputs.

Fig. 4 shows a comparison of mean SNR and ITA° value per volunteer for the two SLSC imaging approaches. SNR generally improved with DNN SLSC imaging relative to CPU SLSC imaging for each volunteer, which is attributed to the smoother coherence functions with the DNN approach. In addition, SNR generally increased with an increase in ITA° values (i.e., with lighter skin tones), which is expected because lighter skin tones generally generate less acoustic clutter than darker skin tones with both CPU and DNN SLSC imaging.

Fig. 5 shows normalized CPU and DNN SLSC images derived from artificial Gaussian noise, zeros, and ones photoacoustic channel data inputs, as defined by Zhang *et al.* [16]. A similar Gaussian noise pattern is displayed in both the

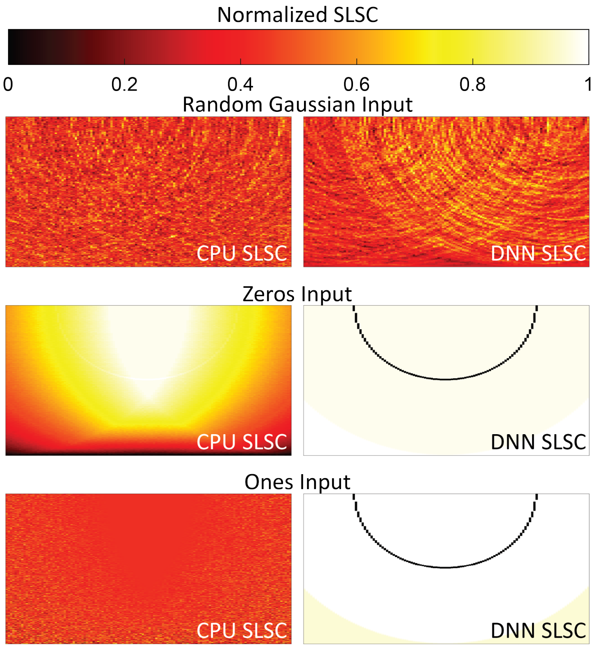

Fig. 5. Normalized CPU SLSC and DNN SLSC images produced with artificial photoacoustic channel data inputs.

CPU and GPU SLSC images which indicates that the DNN has retained the fundamental characteristics of the imaging process. In addition, no vessels are observed in the binary cases, which this is important because the appearance of vessels in these synthetic cases would indicate that the DNN is imposing learned features from the training dataset onto an unrelated input, which is one key hallmark of overfitting in ultrasound beamforming [16], [19]. These observations suggest that our trained DNN did not overfit to the training dataset.

IV. Conclusion

This work is the first to successfully translate CohereNet from ultrasound to photoacoustic data, enabled by the variability introduced with different SNR, clutter, and skin tones included in the photoacoustic training dataset. Notably, the DNN SLSC method reduced acoustic clutter in the background, thereby improving SNR at target vessels. In addition, transfer learning was achieved without evidence of overfitting. With previous applications to a variety of ultrasound data and with this new application to photoacoustic data, the results herein highlight the versatility CohereNet.

Acknowledgment

This work was completed through the visiting undergraduate student program at Johns Hopkins University, in partnership with the Pontificia Universidad Católica del Perú (home institution) and the University of São Paulo (collaborating institution providing *in vivo* data). Partial financial support was derived from the Alfred P. Sloan Foundation, the Chan Zuckerberg Initiative (Grant No. 2022-309513), and the National Institutes of Health (Grant No. R01 EB032960).

References

[1] A. B. E. Attia, G. Balasundaram, M. Moothanchery, U. Dinish, R. Bi, V. Ntziachristos, and M. Olivo, "A review of clinical photoacoustic imaging: Current and future trends," *Photoacoustics*, vol. 16, p. 100144, Dec. 2019.

[2] J. Xia, M. A. L. Bell, J. Laufer, and J. Yao, "Translational photoacoustic imaging for disease diagnosis, monitoring, and surgical guidance: introduction to the feature issue," *Biomedical Optics Express*, vol. 12, no. 7, p. 4115, Jun. 2021.

[3] 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.

[4] B. Pourebrahimi, S. Yoon, D. Dopsa, and M. C. Kolios, "Improving the quality of photoacoustic images using the short-lag spatial coherence imaging technique," in *Photons Plus Ultrasound: Imaging and Sensing 2013*, vol. 8581. SPIE, 2013, pp. 680–685.

[5] 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, pp. 126 011–126 011, 2014.

[6] 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*, vol. 67, no. 10, pp. 2069–2084, 2020.

[7] G. S. P. Fernandes, J. H. Uliana, L. Bachmann, A. A. O. Carneiro, M. A. L. Bell, and T. Z. Pavan, "Impact of skin pigmentation on photoacoustic imaging using linear array transducer: a pilot in vivo study," in *2022 IEEE International Ultrasonics Symposium (IUS)*. IEEE, Oct. 2022.

[8] A. Wiacek, E. Gonzalez, and M. A. L. Bell, "CohereNet: A deep learning architecture for ultrasound spatial correlation estimation and coherence-based beamforming," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 12, pp. 2574–2583, Dec. 2020.

[9] E. A. Gonzalez and M. A. L. Bell, "GPU implementation of photoacoustic short-lag spatial coherence imaging for improved image-guided interventions," *Journal of Biomedical Optics*, vol. 25, no. 07, p. 1, Jul. 2020.

[10] A. Wiacek, E. Gonzalez, N. Dehak, and M. A. L. Bell, "CohereNet: A deep learning approach to coherence-based beamforming," in *2019 IEEE International Ultrasonics Symposium (IUS)*. IEEE, Oct. 2019.

[11] G. Paltauf, R. Nuster, and M. Frenz, "Progress in biomedical photoacoustic imaging instrumentation toward clinical application," *Journal of Applied Physics*, vol. 128, p. 180907, 11 2020.

[12] S. Del Bino, S. Ito, J. Sok, Y. Nakanishi, P. Bastien, K. Wakamatsu, and F. Bernerd, "Chemical analysis of constitutive pigmentation of human epidermis reveals constant eumelanin to pheomelanin ratio," *Pigment Cell & Melanoma Research*, vol. 28, no. 6, pp. 707–717, 2015.

[13] 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*, vol. 58, no. 7, pp. 1377–1388, 2011.

[14] F. Chollet *et al.*, "Keras," 2015. [Online]. Available: https://keras.io

[15] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: http://tensorflow.org/

[16] J. Zhang, A. Wiacek, and M. A. L. Bell, "Binary and random inputs to rapidly identify overfitting of deep neural networks trained to output ultrasound images," in *2022 IEEE International Ultrasonics Symposium (IUS)*. IEEE, Oct. 2022.

[17] A. Chardon, I. Cretois, and C. Hourseau, "Skin colour typology and suntanning pathways," *International Journal of Cosmetic Science*, vol. 13, no. 4, pp. 191–208, 1991.

[18] 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*, vol. 140, no. 1, pp. 3–12, 2020.

[19] D. Hyun, A. Wiacek, S. Goudarzi, S. Rothlübbers, A. Asif, K. Eickel, Y. C. Eldar, J. Huang, M. Mischi, H. Rivaz *et al.*, "Deep learning for ultrasound image formation: Cubdl evaluation framework and open datasets," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency control*, vol. 68, no. 12, pp. 3466–3483, 2021.